A.O. Gel'fond and Yu. V. Linnik

Elementary Methods in the Analytic Theory of Numbers

ELEMENTARY METHODS IN THE ANALYTIC THEORY OF NUMBERS

Elementary Methods in the Analytic Theory of Numbers

BY

A. O. GEL'FOND

AND

YU. V. LINNIK

TRANSLATION BY

D. E. BROWN

TRANSLATION EDITED BY

I. N. SNEDDON

THE M.I.T. PRESS

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

CAMBRIDGE, MASSACHUSETTS

Copyright © 1966
Pergamon Press Ltd.

First edition 1966

This is a translation from the original Russian
Элементарные методы в аналитической теории чисел
(Elementarnye metody v analiticheskoi teoriye chisel)
published in 1962 by Fizmatgiz, Moscow

Library of Congress Catalog Card No. 66–25629

2323/66

CONTENTS

FOREWORD ix

CHAPTER 1. ADDITIVE PROPERTIES OF NUMBERS. SCHNIRELMANN'S METHOD. MANN'S THEOREM. ERDÖS'S THEOREM 1

1. Additive Properties of Sequences 1
2. Mann's Theorem 4
3. Essential Components. Erdös's Theorem 10

CHAPTER 2. AN ELEMENTARY SOLUTION OF WARING'S PROBLEM AND THE HILBERT-KAMKE PROBLEM 13

1. A Note on the Waring and Hilbert–Kamke Problems 13
2. The Fundamental Lemma in the Elementary Solution of the Waring Problem 14
3. Lemmas Concerning Linear Equations 15
4. Proof of the Fundamental Lemma 23
5. Further Estimates for $r_k(m)$ 24
6. Completion of the Proof 28
7. Statement of the Hilbert–Kamke Problem 30
8. Sequences of Integral-valued Vectors and their Densities 31
9. Some Lemmas 32
10. Proof of the Fundamental Lemma 36

CHAPTER 3. THE PROBLEM OF THE DISTRIBUTION OF PRIMES 42

1. Numerical Functions and the Relationships among them. Estimation of the Number of Primes in a Segment of the Series of Positive Integers 42
2. Dirichlet's Theorem on the Infinity of Primes in an Arithmetic Progression 46
3. Fundamental Inequalities for Estimating the Number of Primes in the Series of Positive Integers 50
4. Fundamental Inequalities for Estimating the Number of Primes in a Progression 60
5. Proof of the Limit Theorems for the Distribution of Primes in the Natural Series and in Progressions 64

732274

6. Primes in Sequences of a Rather More General Type than Progressions 69

Chapter 4. An Elementary Derivation of the Law of Distribution of Gaussian Primes. A Theorem on Almost-prime Gaussian Numbers 77

1. Introduction 77
2. Some Auxiliary Formulae 79
3. Proof of the Formula for

$$\sum_{\varrho \in D} \ln^2 |\varrho| + \frac{1}{4} \sum_{\substack{\varrho, \sigma \\ \varrho, \sigma \in D}} \ln |\varrho| \ln |\sigma| \quad 82$$

4. A Recurrence Formula for the Remainder Term 85
5. "Islets" with Small Values of $|G(D/\nu)|$ 90
6. Proof of the Theorem 95
7. A Theorem on Almost-prime Gaussian Numbers 99

Chapter 5. The Sieve of Eratosthenes 101

1. The "Double Rectangular Sieve" 101
2. Viggo Brun's Sieve 105

Chapter 6. Atle Selberg's Method 119

1. Selberg's Asymptotic Formulae 119
2. Schnirelmann's Theorem 127

Chapter 7. The Distribution of the Fractional Parts of Numerical Sequences 130

1. Statement of the Problem 130
2. A Lemma on Differences 131
3. A Lemma on an Incomplete System of Residues 135
4. The Comparison of Two Sums 137
5. Vinogradov's Elementary Derivation of Some Theorems on Sequences of Primes 140
6. Proof of the Theorem 149

Chapter 8. Computation of the Integral Points within Contours 153

1. Statement of the Problem. Typical Problems 153
2. The Statement of Vinogradov's Theorem 156
3. Application of Sonin's Formula 165
4. Generalization of Theorem 8.2.1 167
5. Distribution over a Closed Contour 169

CHAPTER 9. THE DISTRIBUTION OF POWER RESIDUES 171

1. A Theorem of Vinogradov 171
2. Proof of Theorem 9.1.1 176
3. Other Elementary Theorems on the Distribution of Characters. Unsolved Problems 178
4. Elementary Consequences of a Non-elementary Theorem 183

CHAPTER 10. AN ELEMENTARY PROOF OF HASSE'S THEOREM 193

1. Statement of the Problem 193
2. Addition of Solutions 194
3. The Fundamental Construction 197
4. Deduction of the Theorem from the Fundamental Lemma 199
5. Proof of the Fundamental Lemma 200

CHAPTER 11. AN ELEMENTARY PROOF OF SIEGEL'S THEOREM 205

1. Statement of the Theorem. Method of Proof 205
2. Lemmas 206
3. Proof of the First Fundamental Lemma 209
4. Continuation of the Proof. Further Lemmas 211
5. The Second Fundamental Lemma. Completion of the Proof 217

CHAPTER 12. THE TRANSCENDENCE OF CERTAIN NUMBER CLASSES 219

1. Auxiliary Propositions 219
2. General Theorems on the Transcendence of e^{ω} and α^{β} for Algebraic and Real ω, α, β 222

REFERENCES 227

FOREWORD

MANY important problems of the modern analytic theory of numbers can be stated in terms of elementary mathematics and the concept of limit or even simply the concept of an indefinitely increasing parameter. Examples are the prime number theorem, Vinogradov's theorem that all sufficiently large odd numbers are representable as the sum of three primes and that the number of such representations is expressed by a simple asymptotic formula, theorems on computation of points with integral coordinates inside an expanding contour, and on the behaviour of the fractional parts of sequences, etc. Apart from this, the solution of the corresponding problems, stated in simple terms, has often required extremely complicated devices, at first sight remote from the theory of numbers. Until recently, for instance, the prime number theorem could only be proved with the aid of the theory of functions of a complex variable, and the appearance in 1949 of completely elementary proofs by A. Selberg, and by P. Erdős and A. Selberg, was a great event in the theory of numbers. The theory of functions of a complex variable has also played an essential role in additive problems (the original versions of the Hardy–Littlewood solutions of Waring's problem) in various theorems on the distribution of primes and their generalizations, A. O. Gel'fond's solution of Hilbert's seventh problem and many other cases. Fourier series and trigonometric sums play a fundamental part in additive problems, while the theory of fields of functions over an abstract field and algebraic topology are of ever-increasing significance in modern number theory. The transcendental methods mentioned lead in numerous cases to extremely strong and precise results, and much more can be expected from them. There can be no question of renouncing transcendental methods in modern number theory. Yet the investigator feels a natural desire to seek more arithmetic approaches to the solution of problems which can be stated in elementary terms. Apart from the obvious methodological value of such an approach, it is important by reason of giving a simple

and natural insight into the theorems obtained and the causes underlying their existence. It is often possible to attain by elementary methods results which are inaccessible to the powerful analytic methods which operate so successfully in other cases. This is the position, for instance, with binary problems of the Goldbach type; the most important results are achieved here with the aid of the elementary method of the sieve of Eratosthenes, developed by Viggo Brun. In the majority of familiar cases, however, elementary methods, while basically yielding a solution of the problem, still fall short of the transcendental methods in attaining precision in the limit relationships obtained.

At the present time very many problems of the analytic theory of numbers are capable of solution by elementary methods, and solutions have been published in various journals and monographs. The aim of the present volume is to collect these solutions, to systematize them to some extent and to introduce simplifications where possible. We include here elementary methods in additive problems, in problems of computing the integral points inside a contour (asymptotic geometry of numbers), and in the theory of the distribution of primes. Chapter 10 contains an elementary proof by Yu. I. Manin of the Riemann hypothesis for elliptic fields of functions over a finite field (theorems on third power congruences). This proof, originally published in 1956, has been revised and simplified by its author. Chapter 12 discusses the elementary proof originally given by A. O. Gel'fond of this theorem on the transcendence of numbers of the form α^{β} (α, β are algebraic numbers, $\alpha \neq 0$ and 1; β irrational) for the case of real α and β.

Notice that there is a variation of the extent to which the methods discussed in the various chapters are elementary. In problems of the computation of the integral points in contours, it is in essence impossible to avoid the concepts of area enclosed by a curve, of tangent and radius of curvature, i.e. the concepts of definite integral and derivative. These concepts are employed in their simplest form in the relevant chapters. In Chapter 12, "The Transcendence of certain number classes", the concepts of algebraic number and derivative are employed. Those sections in which the derivative or the integral are employed are marked by an asterisk. There are only a few such sections in the book.

The reader requires some knowledge of the elementary theory of numbers, such as may be found in I. M. Vinogradov's *Foundations*

of the Theory of Numbers (Osnovy teorii chisel) or in the first four chapters of B. A. Venkov's *Elementary Theory of Numbers* (Elementarnaya teoriya chisel).

A. O. Gel'fond wrote Chapters 3 and 12, and his co-author Yu. V. Linnik Chapters 1, 2, 4, 7, 8, 9 and 11. In addition, A. I. Vinogradov wrote Chapters 5 and 6, and Yu. I. Manin Chapter 10.

CHAPTER 1

ADDITIVE PROPERTIES OF NUMBERS. SCHNIRELMANN'S METHOD. MANN'S THEOREM. ERDÖS' THEOREM

1. Additive Properties of Sequences

We shall consider infinite sequences of integers, starting with zero:

$$0, a_1, a_2, a_3, \ldots \tag{1.1.1}$$

Here $0 < a_1 < a_2 < \ldots$

We shall denote sequences of this type by Roman capitals; for example, the sequence (1.1.1) will be denoted by A. We shall sometimes have occasion to consider different sequences of the same form; in this case we shall furnish the numbers of the form (1.1.1) with two subscripts:

$$0, a_{i1}, a_{i2}, a_{i3}, \ldots \tag{1.1.2}$$

where $i = 1, 2, \ldots, N$; we thus form N sequences A_i.

Given N sequences A_i ($i = 1, 2, \ldots, N$) (some of which may be identical), we shall describe as the sum of the A_i the new sequence $A_1 + \ldots + A_N$, of the form

$$0, b_1, b_2, b_3, \ldots,$$

where

$$b_j = a_{1j_1} + a_{2j_2} + \ldots + a_{Nj_N}$$

is the sum of any of the N numbers of our sequences, each taken once. Terms of the form a_{i0} ($i = 1, 2, \ldots, N$), which we shall take to mean the number 0, are also permissible here; the numbers b_i obtained will be reckoned once and only once, and will be arranged in increasing order.

The additive theory of numbers investigates the properties of the operation of adding sequences; in particular it describes the numbers resulting from the addition of sequences. Cases

can be found when, as a result of the addition of a bounded number of sequences, all the integers, or all sufficiently large integers, are obtained.

If a sequence A, after being added to itself k times, yields all the integers, it is called a basis of the kth order. (Obviously, it will then also be a basis of order $k_1 > k$.) There is a familiar theorem of Lagrange in elementary number theory which states that every number is the sum of four squares. The sequence Q of squares is therefore a basis of the fourth order. The sequence of cubes forms a ninth order basis (this is already difficult to prove by elementary methods).

The basis properties of sequences are occasionally discussed in the case when only sufficiently large numbers are taken into account. If a sequence A yields after k additions a sequence in which all sufficiently large numbers make an appearance, we shall call it a kth order basis for sufficiently large numbers. This concept is useful since in many cases the fundamental properties of sums of sequences are revealed when considering large numbers; naturally, only a few initial terms of the sequence will have any influence on the structure of its small terms.

According to a familiar theorem of Vinogradov,[1] three times addition of the sequence P of primes (including 0) forms the sequence $P + P + P$ in which all sufficiently large odd numbers are to be found. The sequence P therefore forms a fourth order basis for sufficiently large numbers. The sequence of cubes, which we mentioned above, forms a seventh order basis for sufficiently large numbers.[13], [14] In 1930 L. G. Schnirelmann[30] posed the problem of investigating the addition of sequences of a general type, about which the only information is in regard to the "density" of the distribution of terms. The concept of the "density" of a sequence A is defined by Schnirelmann as follows:

Let $A(n) = \sum_{a_i \leqslant n} 1$ be the numerical function of the sequence, namely the number of its terms not exceeding n (discounting $a_0 = 0$). Then

$$0 \leqslant \frac{A(n)}{n} \leqslant 1.$$

The density $d(A)$ of the sequence A is defined as

$$d(A) = \inf_n \frac{A(n)}{n}. \tag{1.1.3}$$

As an example, notice that, if $1 \notin A$, then $d(A) = 0$. Further, $d(A) = 1$ when and only when A contains the entire set of positive integers. The density of the sequence of squares, cubes or primes is 0. Schnirelmann proved the following important theorem on the density of the sum of two sequences A and B.

THEOREM 1.1.1 (Schnirelmann).

$$d(A + B) \geqslant d(A) + d(B) - d(A)\,d(B). \tag{1.1.4}$$

To prove this, we consider the positive integers of the segment $[1, n]$ and introduce the numerical functions $A(n)$ and $B(n)$ of A and B. Let us also write $A + B = C$; $d(A) = \alpha$; $d(B) = \beta$; $d(C) = \gamma$. The segment $[1, n]$ contains $A(n)$ numbers of A, each of which also appears in C. Let a_k, a_{k+1} be two adjacent numbers of this type. Between them lie the numbers

$$a_k + 1,\ a_k + 2,\ \ldots,\ a_k + l = a_{k+1} - 1,$$

which do not belong to A. Certain of these numbers will belong to $C = A + B$: as for instance the numbers $a_k + r$, where $r \in B$. The number of numbers of the latter type will obviously be $B(l)$. Hence

$$C(n) \geqslant A(n) \sum_{(l)} B(l),$$

where the summation is over the segments between numbers of the type a_k and a_{k+1}. Furthermore, $B(l) \geqslant \beta l$. Hence

$$C(n) \geqslant A(n) + \beta \sum l = A(n) + \beta(n - A(n)) =$$
$$= A(n)\,(1 - \beta) + \beta n \geqslant \alpha(1 - \beta)\,n + \beta n$$

for all $n \geqslant 1$. Hence

$$\gamma = d(C) = \inf_n \frac{C(n)}{n} \geqslant \alpha + \beta - \alpha\beta,$$

which proves (1.1.4).

Inequality (1.1.4) may conveniently be written as

$$1 - d(A + B) \leqslant (1 - d(A))\,(1 - d(B)), \tag{1.1.5}$$

whence it is easy to deduce by induction an inequality which is valid for any number of terms:

$$d(A_1 + \cdots + A_k) \geqslant 1 - \prod_{i=1}^{k} (1 - d(A_i)). \tag{1.1.6}$$

This last inequality leads to a simple but important theorem of Schnirelmann.

THEOREM 1.1.2 (Schnirelmann). *Every sequence of positive density is the basis of the series of positive integers.*

In fact, we have from (1.1.6), for the same sequence A, where

$$\underbrace{A + \cdots + A}_{k \text{ times}} = A^{(k)}$$

$$d(A^{(k)}) \geqslant 1 - (1 - \alpha)^k.$$

With sufficiently large k,

$$d(A^{(k)}) > \tfrac{1}{2}. \tag{1.1.7}$$

In view of this, the number of numbers of $A^{(k)}$ on the segment $[1, n]$ will be greater than $n/2$. If a_i runs over the system of numbers $A^{(k)}$ not exceeding n, there will be altogether more than $n + 1$ numbers $|a_i|$ (including 0) and numbers $|n - a_i|$, and they cannot all be different. Hence we have: $n = a_{i1} + a_{i2}$ is the sum of two numbers of $A^{(k)}$, so that A is a basis of order not exceeding $2k$. This theorem finds important applications in subsequent chapters.

2. Mann's Theorem

Schnirelmann's inequality (1.1.4) is not exact and can be substantially strengthened. As above, let $C = A + B$.

THEOREM 1.2.1 (Mann).

$$d(C) \geqslant \min\,(d(A) + d(B),\ 1)\dagger. \tag{1.2.1}$$

We can assume without loss of generality that $d(A) + d(B) \leqslant 1$; if this is not the case, we can always discard from one of the sequences a certain number of sufficiently remote numbers, so that, for the new sequences A' and B', $d(A') + d(B') < 1$ and $d(A') + d(B')$ is as close as desired to 1. It will then follow from (1.2.1) that $d(C) \geqslant d(A' + B')$ is as close as desired to 1, i.e. $d(C) = 1$.

† This theorem has occupied many mathematicians; it was proved by Mann in 1942;[42] a simpler proof was given by Artin and Scherk[32] in 1949, while the proof was simplified still further by A. Ya. Khinchin,[28] whom we follow here.

We therefore assume that $d(A) + d(B) \leqslant 1$. We have to show that

$$d(C) \geqslant d(A) + d(B). \tag{1.2.2}$$

The proof is based on a fundamental lemma.

LEMMA. *For every* $n \geqslant 1$, *an* $m \in [1, n]$ *can be found such that*

$$C(n) - C(n-m) \geqslant (\alpha + \beta)\, m, \tag{1.2.3}$$

(*as above*, $\alpha = d(A)$; $\beta = d(B)$; $\gamma = d(C)$).

The proof of this lemma is fairly difficult, though inequality (1.2.2) and the fundamental theorem (1.2.1) may be deduced from it without difficulty.

Suppose (1.2.3) has been proved. We consider the segment $[1, n - m]$. By the lemma, a segment $[n - m - m' + 1, n - m]$ can be found here, in which the number of numbers of C is not less than $(\alpha + \beta)\, m'$. We extract from $[1, n - m - m']$ the extreme right-hand segment with the same property then continue our argument. As a result we obviously arrive at the inequality (1.2.2).

We now turn to the proof of the lemma. We shall consider a subset of the segment of numbers $0, 1, \ldots, n$. We shall assume that $n \notin C$ since otherwise the lemma is trivial for the number n: when $m = 1$,

$$C(n) - C(n-m) = C(n) - C(n-1) \geqslant (\alpha + \beta)\, 1.$$

We shall describe the system of numbers $H \subset [0, n]$ as *normal* if, for any numbers $f \in [1, n]$; $f' \in [1, n]$; $f \notin (H)$; $f' \notin H$, we have $f + f' - n \notin H$. If the sequence C possesses the property that segments of it in the segment $[0, n]$ form a normal system, the lemma is easily proved for $C(n)$. In fact, let m be the least positive integer not belonging to C; then $m < n$, since $n \notin C$ by hypothesis.

Let $s \in (n - m, n)$; then $0 < s + m - n < m$. Now $s \in C$. If this were not the case, we should have $s \notin C$; $m \notin C$. In view of the normal property, $s + m - n \notin C$, but $s + m - n < m$, whilst m is the least number not belonging to C. In view of this, all the numbers $s \in (n - m, n)$ belong to C. Consequently,

$$C(n) - C(n-m) = m - 1.$$

Since $m \notin C = A + B$, we must have $A(m) + B(m) \leqslant m - 1$, as may easily be seen by arguing as in the proof of Theorem 1.2.1.

Hence

$$C(n) - C(n-m) \geqslant A(m) + B(m) \geqslant (\alpha + \beta)\, m,$$

which proves (1.2.3).

We must thus now assume that $C = A + B$ does not have the property of normality described above. In this case we introduce a special method of extension of the sets B and C, which we shall term "canonical extension". Suppose C does not possess the property of normality. Then there exist numbers c, c' in the interval $(0, n)$ such that

$$c \notin C; \quad c' \notin C; \quad c + c' - n \in C.$$

Since $C = A + B$, it follows that

$$c + c' - n = a + b \qquad (a \in A,\ b \in B). \tag{1.2.4}$$

Let $\beta_0 \in B$ be the least number for which (1.2.4) is fulfilled. It will be termed the base of our extension. The equation

$$c + c' - n = a + \beta_0$$

will have a solution in the numbers c, c', a, satisfying the conditions: $c \notin C$, $c' \notin C$; $a \in A$; c, c', $a \in (0, n)$. We combine all the numbers c, c' satisfying the above equation and the given conditions into the set C^*. Now $C \cap C^*$ is the empty set. The union

$$C \cup C^* = C_1$$

will in fact be our "canonical extension" of the set C. We consider the expression

$$\beta_0 + n - c. \tag{1.2.5}$$

If c runs over the numbers of the set C^*, the values of $\beta_0 + n - c$ form a set B^*; and in view of the above equation, the numbers (1.2.5) can be written in the form $c' - a$, where $c' \in C^*$, $a \in A$.

Let b^* be any number of B^*. Since it has the form $\beta_0 + n - c$, then $b^* \geqslant \beta_0 \geqslant 0$; since $b^* = c' - a$, then $b^* \leqslant c' \leqslant n$, so that all the numbers of B^* lie in the interval $(0, n)$. Further, since $b^* \in B^*$, then $b^* \notin B$, since otherwise $b^* = c' - a$ would imply $c' = a + b^* \in A + B = C$, which is false. Hence $B^* \cap B$ is the empty set; B^* lies in the segment $(0, n)$. Let us put

$$B \cup B^* = B_1.$$

The set B_1 will be termed the canonical extension of the set B.

We now show that

$$A + B_1 = C_1.$$

Let $a \in A$, $b_1 \in B_1$; we show that $a + b_1 \in C_1$. Either $b_1 \in B$, or $b_1 \in B^*$. If $b_1 \in B$, then $a + b_1 \in A + B = C \subset C_1$; if $b_1 \in B^*$, then $a + b_1 \in C_1$ or $a + b_1 \notin C$. In this case, since $b_1 \in B^*$, $b_1 = \beta_0 + n - c'$; $c' \notin C$, so that

$$c'' = a + b_1 = a + \beta_0 + n - c' \notin C$$

and

$$c'' + c' - n = a + \beta_0 \in A + B = C; \quad c'' \notin C; \quad c' \notin C;$$

by definition of C^*, we find that

$$c'' = a + b_1 \in C^* \subset C_1.$$

Thus $A + B_1 \subset C_1$. Conversely, let $c \in C_1$, so that $c \in C$ or $c \in C^*$. If $c \in C$, then $c = a + b$; $a \in A$; $b \in B \subset B_1$. Whereas if $c \in C^*$, then $b^* = c - a$ belongs to B^* for some a. We have: $c = a + b^* \in A + B^* \subset A + B_1$. Hence

$$C_1 = A + B_1.$$

Further, $n \notin C_1$. For, $n \bar{\in} C$; if we had $n \in C^*$, we should be able to put $c' = n$ in the relationship $c + c' - n = a + \beta_0$, and we should obtain: $c = a + \beta_0 \in A + B = C$, which is impossible, since $c \notin C$.

If the extended sequence C_1 is still not normal, we consider the sets A, B_1, C_1; here $A + B_1 = C_1$; $n \notin C_1$; the position is completely analogous to the previous and we can apply the foregoing arguments to the sets A, B_1, C_1.

We again find the base β_1 of the extensions of the sets B_1^*, C^*, and we put

$$B_1 \cup B_1^* = B_2; \quad C_1 \cup C_1^* = C_2$$

and find that

$$A + B_2 = C_2; \quad n \notin C_2.$$

We continue this process until one of the extensions of sets C_k is normal. Such a point is bound to be reached, since we introduce with each extension new numbers into the interval $(0, n)$. We thus obtain a finite series of sets

$$B = B_0 \subset B_1 \subset \cdots \subset B_k.$$
$$C = C_0 \subset C_1 \subset \cdots \subset C_k.$$

Here, $B_{\mu+1}$ contains the numbers not belonging to B_μ and forming B_μ^*, so that $B_{\mu+1} = B_\mu \cup B_\mu^*$; the same applies for $C_{\mu+1} = C_\mu \cup C_\mu^*$ $(0 \leqslant \mu \leqslant k - 1)$.

Let β_μ be the base of the extension from B_μ and C_μ to $B_{\mu+1}$ and $C_{\mu+1}$. We have

$$A + B_\mu = C_\mu; \quad n \notin C_\mu \qquad (0 \leqslant \mu \leqslant k).$$

Further, C_k is a normal set, while the C_μ with $0 \leqslant \mu \leqslant k-1$ are not normal.

We shall subsequently require certain properties of canonical extensions.

Let us start by proving the inequalities

$$\beta_\mu > \beta_{\mu-1} \qquad (1 \leqslant \mu \leqslant k-1). \tag{1.2.6}$$

Notice that $\beta_\mu \in B_\mu = B_{\mu-1} \cup B^*_{\mu-1}$. If $\beta_\mu \in B^*_{\mu-1}$, then

$$\beta_\mu = \beta_{\mu-1} + n - c,$$

where $c \in C^*_{\mu-1} \subset C_\mu$, so that $c < n$ and $\beta_\mu > \beta_{\mu-1}$. If, on the other hand, $\beta_\mu \in B_{\mu-1}$, then by definition of β_μ there exist numbers $a \in A$, $c \notin C_\mu$, $c' \notin C_\mu$, such that

$$c + c' - n = a + \beta_\mu \in A + B_{\mu-1} = C_{\mu-1}. \tag{1.2.7}$$

In view of this, in accordance with the minimal property of $\beta_{\mu-1}$, $\beta_\mu \geqslant \beta_{\mu-1}$. But if we had $\beta_\mu = \beta_{\mu-1}$, we should have from (1.2.7), by definition of $C^*_{\mu-1}$:

$$c \in C^*_{\mu-1} \subset C_\mu; \quad c' \in C^*_{\mu-1} \subset C_\mu,$$

which is false.

Thus $\beta_\mu \neq \beta_{\mu-1}$, and $\beta_\mu > \beta_{\mu-1}$. We denote by m the least positive number not belonging to C_k. We shall prove the following assertion:

$$\text{If} \quad c \in C^*_\mu \quad (0 \leqslant \mu \leqslant k-1) \quad \text{and} \quad c \in (n-m, n), \quad \text{then}$$
$$c \in (n - m + \beta_\mu, n).$$

It is sufficient to show that $c + m - n > \beta_\mu$. It follows from $c \in (n-m, n)$ that $m + c - n \in (0, m)$ and, by definition of m,

$$m + c - n \in C_k.$$

In view of the fact that

$$C_k = C_\mu \cup C^*_\mu \cup C^*_{\mu+1} \cdots \cup C^*_{k-1},$$

we shall in future distinguish two cases:

(1) $m + c - n \in C_\mu$.

In this case

$$m + c - n = a + b_\mu; \quad a \in A, \quad b_\mu \in B_\mu.$$

But $m \notin C_\mu$ and $c \notin C_\mu$, since $c \in C_\mu^*$. By virtue of the minimality of β_μ, we must have $b_\mu \geqslant \beta_\mu$. But if $b_\mu = \beta_\mu$, we should have $m \in C_\mu^*$ by definition of C_μ^*, which is false, since $C_\mu^* \subset C_\mu \subset C_k$ while $m \notin C_k$. Hence $b_\mu > \beta_\mu$, and consequently

$$m + c - n = a + b_\mu \geqslant b_\mu > \beta_\mu,$$

which proves our assertion.

(2) If $c' = m + c - n \in C_\nu^*$ $(\mu \leqslant \nu \leqslant k-1)$, then, by definition of C_ν^*, we must have

$$c' - a = \beta_\nu + n - c''; \quad a \in A, \quad c'' \in C_\nu^*,$$

whence $c' \geqslant c' - a > \beta_\nu \geqslant \beta_\mu$ (in accordance with (1.2.6)), and our assertion is proved.

We shall prove further the equality

$$C_\mu^*(n) - C_\mu^*(n-m) = B_\mu^*(m-1) \qquad (0 \leqslant \mu \leqslant k-1). \quad (1.2.8)$$

We consider the relationship

$$b = \beta_\mu + n - c; \qquad (1.2.9)$$

by definition of B_μ^* and C_μ^*, it follows from $c \in C_\mu^*$ that $b \in B_\mu^*$ and vice versa; if $c \in (n - m + \beta_\mu, n)$, then $b \in (\beta_\mu, m)$ and vice versa. In view of this,

$$C_\mu^*(n) - C_\mu^*(n - m + \beta_\mu) = B_\mu^*(m-1) - B_\mu^*(\beta_\mu).$$

But in view of the previous assertion, $C_\mu^*(n - m + \beta_\mu) = C_\mu^*(n-m)$; further, every $b \in B_\mu^*$ is expressible in the form (1.2.9), where $c < n$, so that $b > \beta_\mu$ and $B_\mu^*(\beta_\mu) = 0$. Hence $C_\mu^*(n) - C_\mu^*(n-m) = B_\mu^*(m-1)$, which is the same as (1.2.8).

We can now turn to the proof of our lemma. Since C_k is normal, inequality (1.2.3) can be applied to A, B_k, C_k, as has been shown above. We have:

$$C_k(n) - C_k(n-m) \geqslant A(m) + B_k(m), \qquad (1.2.10)$$

where m, as before, is the least positive integer not belonging to C_k; we have $m \notin A$ and $m \notin B_k$, so that $A(m) = A(m-1)$; $B_k(m) = B_k(m-1)$. In view of the definitions

$$C_k = C \cup C^* \cup C_1^* \cup \cdots \cup C_{k-1}^*,$$
$$B_k = B \cup B^* \cup B_1^* \cup \cdots \cup B_{k-1}^*.$$

where the member sets of these unions are disjoint, we find, on denoting $C_0^* = C^*$; $B_0^* = B^*$:

$$C_k(n) - C_k(n-m) = C(n) - C(n-m) + \sum_{\mu=0}^{k-1} \{C_\mu^*(n) - C_\mu^*(n-m)\},$$

$$B_k(m) = B_k(m-1) = B(m-1) + \sum_{\mu=0}^{k-1} B_\mu^*(m-1).$$

Hence we find, using (1.2.10):

$$C(n) - C(n-m) + \sum_{\mu=0}^{k-1} \{C_\mu^*(n) - C_\mu^*(n-m)\}$$

$$\geqslant A(m) + B(m-1) + \sum_{\mu=0}^{k-1} B_\mu^*(m-1).$$

We have from (1.2.8):

$$C_\mu^*(n) - C_\mu^*(n-m) = B_\mu^*(m-1) \qquad (0 \leqslant \mu \leqslant k-1).$$

On discarding the sums of equal terms from the foregoing inequality, we get

$$C(n) - C(n-m) \geqslant A(m) + B(m-1) = A(m) + B(m) \geqslant (\alpha + \beta)\, m,$$

which proves our lemma and hence theorem 1.2.1.

3. Essential Components. Erdös's Theorem

We see from Mann's Theorem that, if A is a sequence of positive density, less than 1, and B is another sequence of positive density, the sum of A and B must have a greater density.

Certain sequences of zero density, as well as sequences of positive density, possess the property of B just mentioned. Those sequences B that possess the property of increasing the density of any sequence A of positive density < 1 when added to A, are called *essential components*.

In 1936 Erdös[38] proved by a completely elementary method the following theorem concerning essential components.

THEOREM 1.3.1 (Erdös). *Every basis is an essential component. To be more precise: if B is a basis of the lth order, and A a sequence of positive density $d(A) = \delta > 0$, then when $\delta < 1$,*

$$d(A+B) \geqslant \delta + \frac{\delta(1-\delta)}{2l}. \tag{1.3.1}$$

We shall prove an even more general theorem: let integers which are $\leqslant n$ and do not belong to A be denoted by $b_1, b_2, \ldots$, and let $b_y \leqslant n < b_{y+1}$; $a_x \leqslant n < a_{x+1}$. We put

$$E = b_1 + b_2 + \cdots + b_y - \frac{y(y+1)}{2}.$$

Obviously, $E \geqslant 0$, since $b_i \geqslant i$.

As it turns out, there exist not less than $x + E/ln$ numbers $\leqslant n$ of the form $a + b$, where $a \in A$; $b \in B$; it will only be necessary for us to use the number $b = 0$ and another $b \in B$.

We first prove a lemma.

LEMMA. *There exists an integer $I > 0$ such that there are at least E/n numbers b_r among the numbers $\leqslant n$ of the form $a_1 + I, a_2 + I, \ldots$*

Proof. The number of solutions of the equation

$$a + v = b$$

in positive integers v, a, $b \leqslant n$ is E. In fact, for a given $b = b_r$ there are $b_r - r$ solutions, since the number of numbers $a < b_r$ is $b_r - r$, and each a yields a solution v. On summing over $r = 1, 2, \ldots, y$, we find that the total number of solutions is

$$E = \sum_{r=1}^{y} (b_r - r).$$

But the number of possible values of v does not exceed n, and at any rate there exists one value of v, say I, for which there are not less than E/n solutions of the equation $a + I = b$; $a, b \leqslant n$.

We turn to the proof of the theorem.

Since B is a basis of the lth order, we have

$$I = b^{(1)} + \cdots + b^{(l)}; \quad b^{(v)} \in B$$

We write μ_s for the number of numbers b in the set $a + b^{(s)}$ $(s = 1, 2, \ldots, l)$. We show that

$$\mu_1 + \mu_2 + \cdots + \mu_l \geqslant \frac{E}{n}.$$

In fact, the set of numbers of the form

$$a + b^{(1)} + b^{(2)}$$

contains not more than $\mu_1 + \mu_2$ numbers b. For, the set $a + b^{(1)}$ contains μ_1 numbers b and certain of the numbers a. If we add $b^{(2)}$ to the numbers of the set $a + b^{(1)}$, than μ_1 numbers b yield not more than μ_1 numbers b; the numbers a yield not more than

μ_2 numbers b. We now take the numbers of the form $a + b^{(1)} + b^{(2)} + b^{(3)}$. They contain not more than $\mu_1 + \mu_2 + \mu_3$ numbers b, as is clear if we apply the same arguments to the sum $a + b^{(1)} + b^{(2)}$ and $b^{(3)}$. Similarly, the set of numbers $a + b^{(1)} + \cdots + b^{(l)} = a + I$ will contain not more than $\mu_1 + \cdots + \mu_l$ numbers b. But the set $a + I$ contains at least E/n numbers b, so that at any rate one of the μ_l, say μ_k, satisfies the inequality

$$\mu_k \geqslant \frac{E}{ln},$$

so that the set $a + b^{(k)}$ contains not less than E/ln numbers $b \leqslant n$. But the set $a + b^{(0)}$, where $b^{(0)} = 0$, contains precisely x numbers a. Hence the sets $a + b^{(i)}$ (including $b^{(i)} = 0$) contain at least $x + E/ln$ distinct integers $\leqslant n$. Furthermore, $d(A) = \delta$; $\delta \in (0, 1)$. We have: if $x = f(n)$, then $f(b_r) \geqslant \delta b_r$, whence $b_r - r = f(b_r) > \delta b_r$; $b_r \geqslant r/(1 - \delta)$. Hence

$$E \geqslant \frac{1 + 2 + \cdots + y}{1 - \delta} - \frac{y(y+1)}{2} \geqslant \frac{\delta}{2(1-\delta)}\, y(y+1).$$

We therefore have, for the number N of numbers $\leqslant n$ of the set $a + b^{(i)}$:

$$N \geqslant x + \frac{\delta}{2(1+\delta)} \frac{y^2}{ln} \qquad (y = n - x).$$

We denote the right-hand side of this inequality by $\Psi(x)$. It is easily seen that $\Psi(x)$ is an increasing function; this may be verified by working out that $\Psi(x + 1) - \Psi(x) > 0$. Hence

$$N \geqslant \Psi(x) \geqslant \Psi(\delta n) = \delta n + \frac{\delta}{2(1-\delta)} \frac{(1-\delta)^2 n^2}{ln}$$

$$= n\left(\delta + \frac{\delta(1-\delta)}{2l}\right),$$

which proves our theorem.

Whatever the $m > 0$, the sequence $\{x^m\}$ $(x = 0, 1, 2, \ldots)$ forms a basis, as we said at the beginning of the present chapter and shall prove in the next chapter. This sequence is therefore an essential component. There may exist essential components which are not bases. An example of such a component was given by Linnik;[12] this example is not elementary, however. An example requiring only elementary arguments has been discovered by Stöhr and Wirsing.[50]

CHAPTER 2

AN ELEMENTARY SOLUTION OF WARING'S PROBLEM AND THE HILBERT-KAMKE PROBLEM

1. A Note on the Waring and Hilbert–Kamke Problems

Waring's problem amounts to an investigation of the basis properties of nth power sequences $\{x^n\}$ $(x = 0, 1, 2, \ldots)$, $n \geqslant 1$. As mentioned in Chapter 1, the sequence of squares $(n = 2)$ forms a fourth order basis. Hilbert showed in 1909 that the sequence of nth powers forms a basis, thereby in essence solving Waring's problem. Certain related problems were only solved much later, however, while there are others which still remain unsolved.

Powerful methods have been invented by Hardy and Littlewood, and by Vinogradov, which have not only enabled a number of problems related to Waring's problem to be solved, but have also enabled great advances to be made in other topics of analytic theory of numbers.

The problem has been investigated of the form of the function $g(n)$—the minimal order of the basis of powers $\{x^n\}$ for all numbers; it was solved in essence in 1936 by the Indian mathematician S. Pillai,[43] with the aid of the methods mentioned above. The problem of the behaviour of the function $G(n)$—the minimal basis of nth powers for large numbers—has also been studied. Here, the strongest result has been obtained by Vinogradov:[1]

$$G(n) < n(3 \ln n + 11). \tag{2.1.1}$$

In addition, an asymptotic formula has been investigated for the number of representations of the number N as a sum of s nth powers:

$$x_1^n + \cdots + x_s^n = N. \tag{2.1.2}$$

All the above work on Waring's problem has been carried out by using deep and powerful analytic methods. To date, only the

basic Waring's problem, i.e. the proof that the nth powers of numbers form a basis, has been successfully solved by elementary methods. An elementary proof of this theorem was obtained by Linnik[16] in 1943. A striking form of the proof has been given by Khinchin[28] and it is this that we shall follow.

A problem analogous to Waring's is that of Hilbert concerning the system of Diophantine equations

$$\left.\begin{aligned} x_1^n + \cdots + x_s^n &= N_n, \\ x_1^{n-1} + \cdots + x_s^{n-1} &= N_{n-1}, \\ \cdots\cdots\cdots\cdots\cdots & \\ x_1 + \cdots + x_s &= N_1, \end{aligned}\right\} \tag{2.1.3}$$

where $x_i \geqslant 0$ are integers. This problem was considered by Kamke in 1921; more rigorous results have been obtained by Mardzhanishvili.[19], [20]

Mardzhanishvili's investigations were continued in 1950 by Emel'yanov,[8] who also gave the elementary solution of the Hilbert–Kamke problem that will be described in the present chapter.

2. The Fundamental Lemma in the Elementary Solution of the Waring Problem

Let A_n be a sequence of nth powers; and let $A_n^{(k)}$ be the sum of k sequences A_n. If we can prove that, for sufficiently large k, $A_n^{(k)}$ has a positive density, then theorem 1.2.1 tells us that A_n is a basis of the series of positive integers. In view of this, we shall be concerned with proving that $d(A_n^{(k)}) > 0$ for $k > k_0(n)$. Let $m > 0$.

We consider the equation

$$x_1^n + \cdots + x_k^n = m \qquad (x_i \geqslant 0). \tag{2.2.1}$$

Let $r_k(m)$ be the number of solutions of (2.2.1). Since we shall often encounter in future the phrase "the number (s) of solutions of the equation(s)", we shall abbreviate it as the N.S.E. We shall regard the number $n \geqslant 2$ as fixed. We now formulate the fundamental lemma.

LEMMA 1. *There exists $k = k(n)$ such that, for any $N \geqslant 1$,*

$$r_k(m) = BN^{k/n-1} \qquad (1 \leqslant m \leqslant N) \tag{2.2.2}$$

where B is independent of n.

We show that it follows from Lemma 1 that $d(A_n^{(k)}) > 0$. We know that

$$r_k(0) + r_k(1) + \cdots + r_k(N) = R_k(N)$$

is the number of solutions of the inequality

$$x_1^n + x_2^n + \cdots + x_k^n \leqslant N. \tag{2.2.3}$$

If x_i satisfies the inequality

$$0 \leqslant x_i \leqslant \left(\frac{N}{k}\right)^{1/n} \qquad (i \leqslant k),$$

the system of numbers $(x_1, \ldots, x_k)$ must satisfy the inequality (2.2.3). Hence inequality (2.2.3) has not less than $(N/k)^{k/n}$ solutions, so that

$$R_k(N) \geqslant \left(\frac{N}{k}\right)^{k/n}. \tag{2.2.4}$$

Now suppose that $d(A_n^{(k)}) = 0$. In this case, whatever the $\varepsilon > 0$ and appropriate N, we obtain

$$A_n^{(k)}(N) < \varepsilon N.$$

The N here can be assumed as large as desired. On applying (2.2.2), we find that

$$R_k(N) = \sum_{m=0}^{N} r_k(m) = r_k(0) + \sum_{m=1}^{N} r_k(m) < 1 + BN^{k/n-1}A_n^{(k)}(N) < 1 + B\varepsilon N^{k/n},$$

which contradicts (2.2.4) with sufficiently small ε. Thus $d(A_n^{(k)}) > 0$. We now proceed to the proof of the lemma.

3. Lemmas Concerning Linear Equations

All future equations will be Diophantine (in integers).

LEMMA 2. *Let a_1, a_2, m in the equation*

$$a_1z_1 + a_2z_2 = m \tag{2.3.1}$$

be integers, and $|a_2| \leqslant |a_1| \leqslant A$; $(a_1, a_2) = 1$. The N.S.E. (2.3.1) under the condition $|z_i| \leqslant A$ $(i = 1, 2)$ does not exceed $3A/|a_1|$.

We can assume that $a_1 > 0$, otherwise z_1 can be replaced by $(-z_1)$ in every solution.

Let (z_1, z_2), (z_1', z_2') be two distinct solutions of equation (2.3.1). We find from the equations

$$a_1 z_1 + a_2 z_2 = m,$$
$$a_1 z_1' + a_2 z_2' = m$$

that

$$a_2(z_2' - z_2) = a_1(-z_1' + z_1).$$

Furthermore, $(a_1, a_2) = 1$, so that $z_2' - z_2 \equiv 0 \pmod{a_1}$ and $|z_2' - z_2| \geqslant a_1$.

The number of solutions (z_1, z_2) under the condition $|z_i| \leqslant A$ $(i = 1, 2)$ does not exceed the number of numbers z_2 in these solutions under the condition $|z_2| \leqslant A$. But the difference between the numbers of the form z_2 is divisible by a_1. If the number of them is t, we have

$$a_1(t-1) \leqslant 2A; \quad t \leqslant \frac{2A}{a_1} + 1 \leqslant \frac{3A}{a_1},$$

which it was required to prove.

LEMMA 3. *Suppose that, in the equation*

$$a_1 z_1 + \cdots + a_l z_l = m \tag{2.3.2}$$

we have: $|a_i| \leqslant A$ $(i \leqslant l)$; $\max |a_i| = H (i \leqslant l)$; $(a_1, \ldots, a_l) = 1$. *Then the N.S.E.* (2.3.2) *under the conditions*

$$|z_i| \leqslant A \qquad (i \leqslant l)$$

does not exceed

$$B_l \frac{A^{l-1}}{H}. \tag{2.3.3}$$

We have already proved this lemma with $l = 2$. Suppose that $l \geqslant 3$ and that the lemma has already been proved for the case of $l - 1$ unknowns. Let $\max |a_i| = |a_l| = H$. We shall distinguish two cases:

(1) $a_1 = a_2 = \ldots = a_{l-1} = 0$. Since $(a_1, a_2, \ldots, a_l) = 1$, so that $|a_l| = H = 1$, equation (2.3.2) in fact has the form $z_l = \pm m$, so that z_l takes not more than one value, while all the other z_i do not exceed in absolute value A, and hence take not more than $2A + 1 \leqslant 3A$ values. Hence in our case the N.S.E. (2.3.3) under the conditions indicated does not exceed

$$(3A)^{l-1} = B_l \frac{A^{l-1}}{H}, \tag{2.3.4}$$

which proves the lemma for case (1).

(2) Suppose that at least one of the $a_1, a_2, \ldots, a_{l-1}$ is non-zero. Then there exists

$$(a_1, a_2, \ldots, a_{l-1}) = \delta .$$

We put

$$H' = \max \frac{|a_i|}{\delta} \qquad (i \leqslant l-1).$$

Let the numbers $z_1, z_2, \ldots, z_l$ satisfy the equation (2.3.2) and the inequalities $|z_i| \leqslant A (i \leqslant l)$. We put

$$\frac{a_1}{\delta} z_1 + \cdots + \frac{a_{l-1}}{\delta} z_{l-1} = m', \tag{2.3.5}$$

so that

$$a_1 z_1 + a_2 z_2 + \cdots + a_{l-1} z_{l-1} = \delta m',$$

and

$$\delta m' + a_l z_l = m. \tag{2.3.6}$$

Here

$$|\delta m'| \leqslant \sum_{i=1}^{l-1} |a_i|\,|z_i| \leqslant l\delta H' A ,$$

so that

$$|m'| \leqslant l H' A . \tag{2.3.7}$$

We have thus replaced the initial equation (2.3.2) by the equations (2.3.5) and (2.3.6). In equation (2.3.6) $\delta \leqslant |a_l|$; $(\delta, a_l) = (a_1, \ldots, a_l) = 1$. The N.S.E. (2.3.6), for which $|m'| \leqslant lH'A$, $|z_l| \leqslant A < lH'A$ does not exceed $3lH'A/|a_l|$ by virtue of Lemma 2. With the given m', equation (2.3.5) has, by virtue of the inductive assumption, $C(l)\, A^{l-2}/H$ solutions in integers z_i, $|z_i| \leqslant A$.

Hence it follows that the number of solutions $(z_1, \ldots, z_l)$ of equation (2.3.2) under the conditions $|z_i| \leqslant A\,(i \leqslant l)$ does not exceed

$$\frac{3lH'A}{|a_l|}\, c(l)\, \frac{A^{l-2}}{H'} = B_l \frac{A^{l-1}}{|a_l|} = B_l \frac{A^{l-1}}{H},$$

which proves Lemma 3.

We now consider the set of all equations of the form

$$a_1 z_1 + a_2 z_2 + \cdots + a_l z_l = 0 , \tag{2.3.8}$$

where $|a_i| \leqslant A\,(i \leqslant l)$. Let A_1 be any number connected with A by the relationships $1 \leqslant A \leqslant A_1 \leqslant B_l A^{l-1}$, $l > 2$. We want to find the upper bound of the sum of the N.S.E. (2.3.8) under the condition that $|z_i| \leqslant A_1 (i \leqslant l)$.

LEMMA 4. *The total N.S.E.* (2.3.8) *does not exceed* $B_l (AA_1)^{l-1}$.

We shall again distinguish three cases:

(1) $a_1 = a_2 = \ldots = a_l = 0$ (this case has to be taken into account). For this case the number of solutions $(z_1, \ldots, z_l)$ obviously does not exceed

$$(2A_1 + 1)^l \leqslant (3A_1)^l = B_l A_1^l = B_l A_1^{l-1} B_l A^{l-1} = B_l (AA_1)^{l-1}.$$

(2) At least one of the coefficients a_i is non-zero, and

$$(a_1, a_2, \cdots, a_l) = \delta = 1.$$

Let

$$H = \max(|a_1|, |a_2|, \cdots, |a_l|); \quad H \in [1, A]$$

in this case. Let m be an integer such that

$$\frac{A}{2^{m+1}} < H \leqslant \frac{A}{2^m}. \tag{2.3.9}$$

For one of the equations of form (2.3.8), where $\delta = 1$ and H satisfies inequalities (2.3.9), the number of solutions $z_i (|z_i| \leqslant A_1)$ does not exceed, by virtue of Lemma 3,

$$B_l \frac{A_1^{l-1}}{H} = B_l \frac{A_1^{l-1}}{\frac{A}{2^{m+1}}} = B_l \frac{A_1^{l-1} 2^m}{A}.$$

Further, it is clear from (2.3.9) that $|a_i| \leqslant A/2^m (i \leqslant l)$. Hence the number of equations, where inequalities of the type (2.3.9) are observed, does not exceed

$$\left(2 \frac{A}{2^m} + 1\right)^l = B_l A^l 2^{-ml}.$$

The sum of the corresponding N.S.E. does not exceed

$$B_l \frac{A_1^{l-1} 2^m}{A} B_l A^l 2^{-ml} = B_l (AA_1)^{l-1} 2^{-(l-1)m}.$$

On summing this expression over all $m \geqslant 0$, we find the total upper bound

$$B_l (AA_1)^{l-1}.$$

(3) At least one of the coefficients a_i is non-zero and $(a_1, a_2, \ldots, a_l) = \delta > 1$.

In this case we replace the equation of the form (2.3.8) by

$$\frac{a_1}{\delta} z_1 + \cdots + \frac{a_l}{\delta} z_l = 0; \quad \left(\frac{a_1}{\delta}, \ldots, \frac{a_l}{\delta}\right) = 1.$$

The number A is replaced by the smaller number A/δ. The condition $A_1 = B_l A^{l-1}$ may be infringed, but it was utilized only in (1), and not in (2), and our future discussion will in fact be based on (2). By virtue of the arguments of (2), with a given δ and $|z_i| \leqslant A_1$, the total N.S.E. does not exceed

$$B_l \left(\frac{A}{\delta} A_1\right)^{l-1} = B_l \frac{(AA_1)^{l-1}}{\delta^{l-1}} . \tag{2.3.10}$$

Further, $\delta \leqslant A$. The total N.S.E. with different δ does not exceed

$$B_l (AA_1)^{l-1} \sum_{\delta=1}^{A} \frac{1}{\delta^{l-1}} . \tag{2.3.11}$$

With $q \geqslant 2$, we can write the elementary inequality

$$\sum_{n=1}^{A} \frac{1}{n^q} < \frac{q}{q-1} .$$

On using this inequality with $q = l - 1 \geqslant 2$, we obtain for (2.3.11) the upper bound

$$B_l (AA_1)^{l-1}. \tag{2.3.12}$$

In conjunction with the result of Case (1), (2.3.12) proves Lemma 4.

We shall now consider finite sets of integers in which repetitions are allowed; if the number a appears λ times in the set A, we shall say that its multiplicity in A is $\lambda = \lambda(a)$.

Let c be any given integer. We consider the equations

$$x + y = c; \quad x \in A; \quad y \in B. \tag{2.3.13}$$

LEMMA 5. *In equation* (2.3.13) *the number of solutions does not exceed half the sum of the numbers of solutions of*

$$x - y = 0 \qquad (x \in A, \, y \in A)$$

and

$$x - y = 0 \qquad (x \in B, \, y \in B).$$

We remark for the proof that each value $x \in A$ in (2.3.13) corresponds to not more than one value $y = y_x \in B$, while if we take account of multiplicity, for a given x we obtain not more than $\lambda(x)\,\mu(y_x)$ values, where $\lambda(x)$ is the multiplicity of $x \in A$ and $\mu(y_x)$ the multiplicity of $y_x \in B$. Further, we obviously have $\lambda(x)\,\mu(y_x) \leqslant \frac{1}{2}((\lambda(x))^2 + (\mu(y_x))^2)$. On summing over x and remembering that distinct values of x must correspond to distinct values of y_x, we find that the total upper bound of the N.S.E. (2.3.13)

does not exceed

$$\frac{1}{2}\sum_{x\in A}(\lambda(x))^2+\frac{1}{2}\sum_{y\in B}(\mu(y))^2, \tag{2.3.14}$$

which is obviously the same as half the sum of the numbers of solutions of the equations given above.

An immediate corollary of the lemma is that the same result holds when $A = B$.

LEMMA 6. *The N.S.E.* $x + y = c (x \in A, y \in A)$ *does not exceed the N.S.E.* $x - y = 0 (x \in A, y \in A)$.

Now let k and s be arbitrary positive integers. We put $k\,2^s = l$ and consider the equation

$$x_1 + x_2 + \cdots + x_l = c.$$

Let $A_1, A_2, \ldots, A_l$ be finite sets of numbers; let the set A_i $(1 \leqslant i \leqslant l)$ contain distinct numbers $a_{i1}, a_{i2}, \ldots$, with multiplicities $\lambda_{i1}, \lambda_{i2}, \ldots$ Our interest is in the N.S.E.

$$x_1 + x_2 + \cdots + x_l = c, \quad x_i \in A_i (1 \leqslant i \leqslant l). \tag{2.3.15}$$

If we put

$$x_1 + x_2 + \cdots + x_{l/2} = x, \quad x_{l/2+1} + \cdots + x_l = y$$

($\frac{1}{2}l$ is obviously an integer), the given equation can be rewritten as

$$x + y = c$$

and we can apply to it the Lemma 5 just proved. Let us consider the sets to which the numbers x and y must belong. Since $x_i \in A_i$ $(1 \leqslant i \leqslant l)$, x may be any number of the form $z_1 + z_2 + \ldots + z_{l/2}$, where $z_i \in A_i$ $(1 \leqslant i \leqslant \frac{1}{2}l)$; similarly, y may be any number of the same form, where now $z_i \in A_{l/2+i}$ $(1 \leqslant i \leqslant \frac{1}{2}l)$.

By Lemma 5, therefore, the N.S.E. (2.3.15) does not exceed half the sum of the N.S.E.

$$x - y = 0 \tag{2.3.16}$$

with the following two assumptions:

(1)

$$x = z_1 + z_2 + \cdots + z_{l/2},$$
$$y = z_1' + z_2' + \cdots + z_{l/2}',$$

where

$$z_i \in A_i, \quad z_i' \in A_i \left(1 \leqslant i \leqslant \frac{l}{2}\right). \tag{2.3.17}$$

(2) x and y have the same form, but

$$z_i \in A_{l/2+i}, \quad z_i' \in A_{l/2+i}\left(1 \leqslant i \leqslant \frac{l}{2}\right). \tag{2.3.18}$$

In both cases equation (2.3.16) can be rewritten as

$$(z_1 - z_1') + (z_2 - z_2') + \cdots + (z_{l/2} - z_{l/2}') = 0. \tag{2.3.19}$$

We thus arrive at the result that the N.S.E. (2.3.15) does not exceed half the sum of the N.S.E. (2.3.17) under the assumptions (2.3.17) and (2.3.18), i.e. does not exceed half the sum of the numbers of solutions of

$$\sum_{i=1}^{l/2} (z_i - z_i') = 0, \quad z_i \in A_i, \quad z_i' \in A_i\left(1 \leqslant i \leqslant \frac{l}{2}\right)$$

and

$$\sum_{i=1}^{l/2} (z_i - z_i') = 0, \quad z_i \in A_{l/2+i}, \quad z_i' \in A_{l/2+i}\left(1 \leqslant i \leqslant \frac{l}{2}\right).$$

Equation (2.3.19) has $\frac{1}{2}l$ terms on the left-hand side, i.e. half as many as the initial equation (2.3.15). On putting

$$\sum_{i=1}^{l/4} (z_i - z_i') = x, \qquad \sum_{i=l/4+1}^{l/2} (z_i - z_i') = y,$$

(2.3.19) reduces to the form

$$x + y = 0$$

and we can again apply Lemma 5 to it. In precisely the same way as we arrived from (2.3.15) at the equation (2.3.19), we now evidently arrive from (2.3.19) at the equation

$$\sum_{i=1}^{l/4} (u_i + u_i' - u_i'' - u_i''') = 0, \tag{2.3.20}$$

where we now have to consider the sum of the numbers of solutions of this equation under the following (now four) assumptions:

$$(1)\ u_i, u_i', u_i'', u_i''' \in A_i,$$
$$(2)\ u_i, u_i', u_i'', u_i''' \in A_{l/4+i},$$
$$(3)\ u_i, u_i', u_i'', u_i''' \in A_{l/2+i},$$
$$(4)\ u_i, u_i', u_i'', u_i''' \in A_{3l/4+i}$$
$$\left(1 \leqslant i \leqslant \frac{l}{4}\right).$$

Since $l = k\,2^s$, we can repeat this process s times. As a result, we obviously arrive at the equation

$$\sum_{i=1}^{k} \{y_i^{(1)} + y_i^{(2)} + \cdots + y_i^{(2^{s-1})} - y_i^{(2^{s-1}+1)} - \cdots - y_i^{(2^s)}\} = 0, \quad (2.3.21)$$

where it is now a question of considering the sum of the numbers of solutions of this equation under 2^s different assumptions, namely:

$$\left.\begin{array}{l} (1)\ y_1^{(j)} \in A_1,\ y_2^{(j)} \in A_2,\ \ldots,\ y_k^{(j)} \in A_k, \\ (2)\ y_1^{(j)} \in A_{k+1},\ y_2^{(j)} \in A_{k+2},\ \ldots,\ y_k^{(j)} \in A_{2k}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (2^s)\ y_1^{(j)} \in A_{k2^s-k+1},\ \ldots,\ y_k^{(j)} \in A_{k2^s} \end{array}\right\} (1 \leqslant j \leqslant 2^s)$$

If we put

$$y^{(j)} = y_1^{(j)} + y_2^{(j)} + \cdots + y_k^{(j)} \qquad (1 \leqslant j \leqslant 2^s),$$

equation (2.3.21) takes the simple form

$$y^{(1)} = y^{(2)} + \cdots + y^{(2^{s-1})} - y^{(2^{s-1}+1)} - \cdots - y^{(2^s)} = 0. \quad (2.3.22)$$

We are speaking here of the sum of the numbers of solutions of equation (2.3.22) under the following 2^s assumptions, which differ from one another in the values of the parameter m $(0 \leqslant m \leqslant 2^s - 1)$:

$$y^{(j)} = y_1^{(j)} + y_2^{(j)} + \cdots + y_k^{(j)},$$

where

$$y_1^{(j)} \in A_{mk+1},\ y_2^{(j)} \in A_{mk+2}, \ldots,\ y_k^{(j)} \in A_{(m+1)k}$$
$$(j = 1, 2, \ldots, 2^s).$$

We can therefore formulate our final result as the following lemma.

LEMMA 7. *The N.S.E.*

$$x_1 + x_2 + \cdots + x_l = c \qquad (x_i \in A_i, \quad 1 \leqslant i \leqslant l)$$

does not exceed the sum of the numbers of solutions of

$$y^{(1)} + y^{(2)} + \cdots + y^{(2^{s-1})} - y^{(2^{s-1}+1)} - \cdots - y^{(2^s)} = 0.$$
$$\left.\begin{array}{l} y^{(j)} = y_1^{(j)} + y_2^{(j)} + \cdots + y_k^{(j)}, \\ y_1^{(j)} \in A_{mk+1},\ y_2^{(j)} \in A_{mk+2},\ \cdots,\ y_k^{(j)} \in A_{(m+1)k} \end{array}\right\} (j = 1, 2, \cdots, 2^s)$$

under the assumptions $m = 0, 1, \ldots, 2^s - 1$. *Here* $l = k\,2^s$.

It can be seen that Lemma 5 is a particular case of Lemma 7 with $k = s = 1$, $l = 2$.

4. Proof of the Fundamental Lemma

We turn to the proof of the fundamental Lemma 1. We shall prove it by induction from $n-1$ to n. In order to pass from $n-1$ to n without difficulty, instead of the initial equation

$$x_1^n + x_2^n + \cdots + x_k^n = m \quad (m \leqslant N),$$

of Waring's problem, where $x_i \geqslant 0$, so that $x_i \leqslant m^{1/n} \leqslant N^{1/n} (i \leqslant k)$, we shall consider the equation

$$f(x_1) + f(x_2) + \cdots + f(x_k) = m \; (m \leqslant N), \tag{2.4.1}$$

where $f(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$ is a polynomial which may depend on n and N in such a way that

$$|a_i| = B_n N^{i/n} (0 \leqslant i \leqslant n); \qquad a_0 \neq 0. \tag{2.4.2}$$

A particular case of this polynomial will be $f(x) = x^n$.

The fundamental Lemma 1 will be a consequence of the following more general lemma.

LEMMA 8. *With* $|x_i| \leqslant N^{1/n}$ *and a suitable choice of* $k = k(n)$, *the N.S.E.* (2.4.1) *does not exceed*

$$B_n N^{k/n-1}. \tag{2.4.3}$$

We shall establish this result with $n = 1$; $f(x) = a_0 x + a_1$. We put $k(1) = 2$, so that equation (2.4.1) becomes

$$a_0(x_1 + x_2) = m - 2a_1, \; |x_i| \leqslant N \qquad (i = 1, 2). \tag{2.4.4}$$

Not more than $2N + 1 \leqslant 3N$ values are possible for each x_i, and not more than one x_2 can be found for a given x_1, so that the N.S.E. (2.4.3) does not exceed $3N$, which proves our lemma for $n = 1$. Now let $n > 1$ and suppose the lemma true for $n - 1$. We put $k(n-1) = k'$:

$$k = k(n) = 2n\, 2^{s+1}, \quad s = [4\ln_2 k'] - 1. \tag{2.4.5}$$

We apply Lemma 6 to equation (2.4.1). We put

$$x = \sum_{i=1}^{k/2} f(x_i), \quad y = \sum_{i=k/2+1}^{k} f(x_i). \tag{2.4.6}$$

By Lemma 6, $r_k(m)$ does not exceed the N.S.E. $x - y = 0$, where x, y are given by (2.4.6); $|x_i| \leqslant N^{1/n}$; $|y_i| \leqslant N^{1/n}$ $(1 \leqslant i \leqslant \frac{1}{2}k)$, i.e. the N.S.E.

$$\sum_{i=1}^{k/2} \{f(x_i) - f(y_i)\} = 0 \tag{2.4.7}$$

with the restrictions on $|x_i|$, $|y_i|$ given above.

We carry out a change of variables in (2.4.7): we put $x_i = y_i + h_i$ and consider the system of numbers (y_i, h_i) $(i \leqslant \frac{1}{2}k)$, where y_i, h_i can take arbitrary values in the interval $[-2N^{1/n}, 2N^{1/n}]$, which can only lead to an increase in the N.S.E. (2.4.7). Each term in (2.4.7) transforms to

$$f(y_i + h_i) - f(y_i) = \sum_{v=0}^{n} a_v\{(y_i + h_i)^{n-v} - y_i^{n-v}\} = \sum_{v=0}^{n} a_v \sum_{t=l}^{n=v} C_{n-v}^{t} h_i^t y_i^{n-v-t}.$$

We introduce a new variable by setting

$$v + t = u,$$

so that

$$n - v - t = n - u; \quad t = u - v.$$

We obtain after elementary transformations

$$f(y_i + h_i) - f(y_i) = h_i \varphi_i(y_i),$$

where

$$\varphi_i(y) = \sum_{u=1}^{n} a_{i,u} y^{n-u}$$

is a polynomial of degree $(n - 1)$ with coefficients

$$a_{i,u} = \sum_{v=0}^{n-1} a_v C_{n-v}^{u-v} h_i^{u-v-1} \qquad \left(i \leqslant \frac{k}{2}\right).$$

We see that these coefficients depend on the numbers h_i. The equation (2.4.7) becomes in the new variables (y_i, h_i) $(i \leqslant \frac{1}{2}k)$:

$$h_1\varphi_1(y_1) + h_2\varphi_2(y_2) + \cdots + h_{k/2}\varphi_{k/2}(y_{k/2}) = 0. \qquad (2.4.8)$$

Here $y_i, h_i \in [-2N^{1/n}, 2N^{1/n}]$ and the coefficients $\varphi_i(y_i)$ depend on the h_i.

Thus, $r_k(m)$ does not exceed the total N.S.E. (2.4.8) under the restrictions on $|y_i|$, $|h_i|$ given above.

5. Further Estimates for $r_k(m)$

We have reduced the investigation of $r_k(m)$ to finding the upper bound of the N.S.E. (2.4.8). We shall regard $h_1, \ldots, h_{\frac{1}{2}k}$ as fixed and find a bound under these conditions for the N.S.E. (2.4.8). We shall apply Lemma 7 here, where the role of the x_i will be played by the numbers $h_i\varphi_i(y_i)$, and the role of l by $\frac{1}{2}k = 2n\,2^s$; $k_0 = 2n$.

The sets A_i consist of the numbers $x_i = h_i\varphi_i(y_i)$, where the h_i are given and $|y_i| \leqslant 2N^{1/n}$.

By Lemma 7, under these conditions the N.S.E. (2.4.8) does not exceed the sum of the N.S.E.

$$y^{(1)} + y^{(2)} + \cdots + y^{(2^{s-1})} - y^{(2^{s-1}+1)} - \cdots - y^{(2^s)} = 0 \quad (2.5.1)$$

with 2^s possible conditions, corresponding to the values of the parameter $\mu = 0, 1, \ldots, 2^s - 1$:

$$y^{(j)} = y_1^{(j)} + \cdots + y_{k_0}^{(j)} \qquad (1 \leqslant j \leqslant 2^s),$$
$$y_i^{(j)} \in A_{\mu k_0 + i} \qquad (1 \leqslant i \leqslant k_0).$$

Taking say $\mu = 0$, (2.5.1) will run as follows:

$$\begin{aligned}&\{y_1^{(1)} + y_2^{(1)} + \cdots + y_{k_0}^{(1)}\} + \{y_1^{(2)} + y_2^{(2)} + \cdots + y_{k_0}^{(2)}\}\\ &\quad + \cdots + \{y_1^{(2^{s-1})} + y_2^{(2^{s-1})} + \cdots + y_{k_0}^{(2^{s-1})}\}\\ &\quad - \{y_1^{(2^{s-1}+1)} + y_2^{(2^{s-1}+1)} + \cdots + y^{(2^{s-1}+1)}\}\\ &\quad - \cdots - \{y_1^{(2^s)} + y_2^{(2^s)} + \cdots + y_{k_0}^{(2^s)}\} = 0,\end{aligned}$$

or, on changing the order of the terms:

$$\begin{aligned}&\{y_1^{(1)} + y_1^{(2)} + \cdots + y_1^{(2^{s-1})} - y_1^{(2^{s-1}+1)} - \cdots - y_1^{(2^s)}\}\\ &\quad + \{y_2^{(1)} + y_2^{(2)} + \cdots + y_2^{(2^{s-1})} - y_2^{(2^{s-1}+1)} - \cdots - y_2^{(2^s)}\}\\ &\quad + \cdots + \{y_{k_0}^{(1)} + y_{k_0}^{(2)} + \cdots + y_{k_0}^{(2^{s-1})} - y_{k_0}^{(2^{s-1}+1)} - \cdots - y_{k_0}^{(2^s)}\} = 0;\end{aligned}$$

each of the $y_i^{(j)}$ is here a number of the form $h_i\varphi_i(v_i^{(j)})$, where $|v_i^{(j)}| \leqslant 2N^{1/n}$; the last equation can therefore be rewritten as

$$\begin{aligned}&h_1\{\varphi_1(v_1^{(1)}) + \varphi_1(v_1^{(2)}) + \cdots + \varphi_1(v_1^{(2^{s-1})})\\ &\qquad - \varphi_1(v_1^{(2^{s-1}+1)}) - \cdots - \varphi_1(v_1^{(2^s)})\}\\ &\qquad + h_2\{\varphi_2(v_2^{(1)}) + \cdots - \varphi_2(v_2^{(2^s)})\}\\ &\qquad + \cdots + h_{k_0}\{\varphi_{k_0}(v_{k_0}^{(1)}) + \cdots - \varphi_{k_0}(v_{k_0}^{(2^s)})\} = 0.\end{aligned}$$

On putting for brevity

$$\begin{aligned}&\varphi_i(v_i^{(1)}) + \varphi_i(v_i^{(2)}) + \cdots + \varphi_i(v_i^{(2^{s-1})}) - \varphi_i(v_i^{(2^{s-1}+1)})\\ &\qquad - \cdots - \varphi_i(v_i^{(2^s)}) = z_i \qquad (1 \leqslant i \leqslant k_0)\end{aligned}$$

we can rewrite this equation as

$$h_1 z_1 + h_2 z_2 + \cdots + h_{k_0} z_{k_0} = 0. \quad (2.5.2)$$

We shall always have 2^s equations of this type, and the set of them can be written as

$$\sum_{i=1}^{k_0} h_{mk_0+i} z_{mk_0+i} = 0 \qquad (0 \leqslant m \leqslant 2^s - 1). \quad (2.5.3)$$

We shall only consider equation (2.5.2), since the rest behave similarly. As defined earlier,

$$\varphi_i(y) = \sum_{u=1}^{n} a_{i,u} y^{n-u},$$

where

$$a_{i,u} = \sum_{v=0}^{n-1} a_v C_{n-v}^{u-v} h_i^{u-v-1} \qquad \left(i \leqslant \frac{k}{2}\right).$$

Since we have assumed that $|a_v| \leqslant B_n N^{v/n}$ and $|h_i| \leqslant 2N^{1/n}$, we have

$$|a_{i,u}| = B_n \sum_{v=0}^{n-1} N^{v/n} C_{n-v}^{u-v} N^{\frac{u-v-1}{n}} = B_n N^{\frac{u-1}{n}} \sum_{v=0}^{n-1} C_{n-v}^{u-v},$$

i.e. since $u \leqslant n$,

$$|a_{i,u}| = B_n N^{\frac{u-1}{n}}. \tag{2.5.4}$$

Further, since $|v_i^{(j)}| \leqslant 2N^{1/n}$, we have $|v_i^{(j)}|^{n-u} = B_n N^{(n-u)/n}$, in view of which:

$$|a_{i,u}|\,|v_i^{(j)}|^{n-u} = B_n N^{\frac{n-1}{n}} N^{\frac{n-u}{n}} = B_n N^{\frac{n-1}{n}},$$

whence

$$|\varphi_i(v_i^{(j)})| = B_n N^{\frac{n-1}{n}} \qquad (i \leqslant k_0 2^s;\ j \leqslant 2^s)$$

and

$$|z_i| \leqslant B_n N^{\frac{n-1}{n}} \qquad (i \leqslant k_0 2^s). \tag{2.5.5}$$

Hence, in equation (2.5.3) each value of z_i can only take the bounded values (2.5.5). The z_i can have repetitions. Let us consider the number of such repetitions, i.e. the N.S.E. $z_i = z$, where z is a number satisfying (2.5.5). We have the equation

$$\varphi_i(v_i^{(1)}) + \cdots + \varphi_i(v_i^{(2^{s-1})}) - \varphi_i(v_i^{(2^{s-1}+1)}) - \cdots - \varphi_i(v_i^{(2^s)}) = z, \tag{2.5.6}$$

We shall use here the method of mathematical induction described earlier.

We rewrite the equation as

$$\varphi_i(v_i^{(1)}) + \varphi_i(v_i^{(2)}) + \cdots + \varphi_i(v_i^{(k')}) = z - \varphi_i(v_i^{(k'+1)}) - \cdots + \varphi_i(v_i^{(2^{s-1}+1)}) + \cdots + \varphi_i(v_i^{(2^s)}), \tag{2.5.7}$$

which is possible, since $k' = k(n-1) > 1$, $k(1) = 2$ and we have $2^{s-1} > k'$, as may easily be discovered.

We denote the right-hand side of equations (2.5.7) by m', so that

$$\varphi_i(v_i^{(1)}) + \cdots + \varphi_i(v_i^{(k')}) = m'. \tag{2.5.8}$$

We select for the arguments on the right-hand side of (2.5.7), i.e. for the numbers

$$v_i^{(j)} \qquad (k'+1 \leqslant j \leqslant 2^s),$$

any values from the permissible interval $[-2N^{1/n}, 2N^{1/n}]$; in which case, m' also takes a definite value. We apply to equation (2.5.8) the inductive assumption that we stated earlier. We have:

$$\varphi_i(y) = \sum_{u=1}^{n} a_{i,u} y^{n-u} \quad \text{is a polynomial of degree} \quad n-1;$$

$$|a_{i,u}| = B_n N^{\frac{u-1}{n}} = B_n \left(N^{\frac{n-1}{n}}\right)^{\frac{u-1}{n-1}}, \tag{2.5.9}$$

$|m'| = B_n N^{(n-1)/n}$. In view of the last inequality, the role of N can now be played by the number $c_n N^{(n-1)/n}$ ($c_n > 0$ is a constant) and the necessary conditions are satisfied. In view of this, we can assert that the N.S.E. (2.5.8) with $|v_i^{(j)}| \leqslant 2N^{1/n} = 2(N^{(n-1)/n})^{1/(n-1)}$ has the upper bound:

$$B_n \left(N^{\frac{n-1}{n}}\right)^{\frac{k'}{n-1}-1} = B_n N^{\frac{k'-n+1}{n}}. \tag{2.5.10}$$

The $v_i^{(j)}$ have been fixed here with $k'+1 \leqslant j \leqslant 2^s$. The number of values of these numbers does not exceed

$$(2N^{1/n}+1)^{2^s-k'} = B_n N^{\frac{2^s-k'}{n}}. \tag{2.5.11}$$

The total N.S.E. (2.5.7) has an upper bound obtained by cross-multiplying (2.5.11) and (2.5.10), which gives

$$B_n N^{\frac{2^s-n+1}{n}}. \tag{2.5.12}$$

Hence the multiplicity of each z_i in equation (2.5.2) does not exceed (2.5.12).

It amounts in future to calculating the numbers of solutions of linear equations of the form (2.5.2), where the z_i have multiplicities not exceeding (2.5.12). The numbers h_i are fixed; but we have to look over all the possibilities for them. We have arrived

at the result: the $r_k(m)$ do not exceed the sums of the numbers of solutions in integers z_i, $|z_i| \leqslant B_n N^{(n-1)/n}$ with multiplicities $\lambda_i \leqslant c(n)\, N^{(2^s-n+1)/n}$ of equations of the form

$$\sum_{i=1}^{k_0} h_{\mu k_0+i} z_{\mu k_0+i} = 0, \tag{2.5.13}$$

where μ runs over the values $0, 1, \ldots, 2^s - 1$, while the numbers $h_r (1 \leqslant r \leqslant 2^s k_0)$ run independently of one another over all integers of the interval $(-2N^{1/n}, +2N^{1/n})$.

6. Completion of the Proof

It remains to find the upper bound of the total N.S.E. (2.5.13). Let us first fix the numbers h_i.

We denote by Γ any definite system of h_i ($|h_i| \leqslant 2N^{1/n}$, $1 \leqslant i \leqslant \frac{1}{2}k$) and by $U_\mu(\Gamma)$ the number of solutions of equation (2.5.13) with this fixed system Γ and a given μ, having in mind the solutions z_i satisfying $|z_i| \leqslant B_n N^{(n-1)/n}$ where each z_i is repeated with a multiplicity $\lambda_i \leqslant B_n N^{(2^s-n+1)/n}$. From the result obtained in § 5, we have

$$r_k(m) \leqslant \sum_{\Gamma} \left(\sum_{\mu=0}^{2^s-1} U_\mu(\Gamma) \right),$$

where the summation over Γ is over all permissible systems of numbers h_i. Consequently,

$$r_k(m) \leqslant \sum_{\mu=0}^{2^s-1} \left(\sum_{\Gamma} U_\mu(\Gamma) \right).$$

There is no difference whatever between the equations of the form (2.5.13) with different values of μ; and the relevant restrictions on the solutions are also the same. In view of this, we can concentrate on the equation corresponding to $\mu = 0$, and on the system of numbers $U_0(\Gamma)$. We obtain

$$r_k(m) \leqslant 2^s \sum_{\Gamma} U_0(\Gamma) = B_n \sum_{\Gamma} U_0(\Gamma),$$

where $U_0(\Gamma)$ is the N.S.E.

$$h_1 z_1 + h_2 z_2 + \cdots + h_{k0} z_{k0} = 0 \tag{2.6.1}$$

for the given system Γ of h_i ($|h_i| \leqslant 2N^{1/n}$; $i \leqslant \frac{1}{2}k$). Here $|z_i| \leqslant B_n N^{(n-1)/n}$ and z_i has a multiplicity $\lambda_i \leqslant B_n N^{(2^s-n+1)/n}$. Let

$U_0^*(\Gamma)$ be the N.S.E. (2.6.1) on the assumption that each value of z_i is counted once; obviously,

$$U_0(\Gamma) \leqslant \left(B_n N^{\frac{2^s-n+1}{n}}\right)^{k_0} U_0^*(\Gamma)$$

or, since $k_0 = 2n$,

$$\begin{aligned} U_0(\Gamma) &= B_n N^{2(2^s-n+1)}\, U_0^*(\Gamma), \\ r_k(m) &= B N^{2(2^s-n+1)} \sum_{\Gamma} U_0^*(\Gamma). \end{aligned} \tag{2.6.2}$$

Let us turn to the number system Γ. Each Γ represents a permissible system of values of all the h_i $(1 \leqslant i \leqslant \frac{1}{2}k)$; at the same time, the number $U_0^*(\Gamma)$ is fully defined by the values of the first $k_0 = 2n$ of these numbers $(1 \leqslant i \leqslant 2n)$, since they alone enter into the equation (2.6.1). Having selected some definite system Γ, we thereby uniquely determine some system Γ' of values of the numbers $h_1, h_2, \ldots, h_{2n}$. On the other hand, if a definite system Γ' of numbers $h_1, h_2, \ldots, h_{2n}$ is chosen, it corresponds, not to one system Γ, but to as many as there are methods of "completing the choice" of the remaining h_i $(2n < i \leqslant \frac{1}{2}k)$; since each h_i must belong to the interval $(-2N^{1/n}, +2N^{1/n})$, it is obvious that, for one system Γ', there are not more than

$$c(n)\,(N^{1/n})^{(k/2)-2n} = c(n)\, N^{(k/2n)-2}$$

corresponding systems Γ. Hence

$$\sum_{\Gamma} U_0^*(\Gamma) \leqslant c(n)\, N^{(k/2n)-2} \sum_{\Gamma'} U_0^*(\Gamma').$$

Here $U_0^*(\Gamma)$ is the number of solutions in integers z_i, $|z_i| \leqslant B_n N^{(n-1)n}$ $(1 \leqslant i \leqslant 2n)$ of equation (2.6.1) with the given system Γ' of h_i, $|h_i| \leqslant 2N^{1/n}$ $(1 \leqslant i \leqslant 2n)$, and the summation is carried out over all such systems. We therefore obtain from (2.6.2):

$$\begin{aligned} r_k(m) &= B_n N^{2(2^s-n+1)} N^{(k/2n)-2} \sum_{\Gamma'} U_0^*(\Gamma') \\ &= B_n N^{2(2^{s+1}-n)} \sum_{\Gamma'} U_0^*(\Gamma'). \end{aligned} \tag{2.6.3}$$

We can further apply Lemma 4 (the conditions for it to be applicable are fulfilled). We find that

$$\sum_{\Gamma'} U_0^*(\Gamma') = B_n (AB)^{2n-1} = B_n N^{2n-1}.$$

After this, (2.6.3) gives us

$$r_k(m) = B_n N^{2(2^{s+1}-n)} N^{2n-1} = B_n N^{2\cdot 2^{s+1}-1} = B_n N^{(k/n)-1},$$

which completes the proof of the fundamental Lemma 1 and achieves our purpose.

7. Statement of the Hilbert–Kamke Problem

As mentioned in § 1, this problem consists in investigating the system of diophantine equations

$$\left.\begin{aligned} x_1^n + x_2^n + \cdots + x_S^n &= N_n, \\ x_1^{n-1} + x_2^{n-1} + \cdots + x_S^{n-1} &= N_{n-1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ x_1 + x_2 + \cdots x_S &= N_1. \end{aligned}\right\} \qquad (2.7.1)$$

Here the numbers $x_i \geqslant 0$; the right-hand sides $N_1, N_2, \ldots, N_n$ must be large numbers of different growth orders in correspondence with the left-hand sides. We assume that there exist systems of real numbers $l_1 > 1, \ldots, l_{n-1} > 1$; $i_1 < 1, \ldots, i_{n-1} < 1$ such that

$$l_k N_n^{k/n} \leqslant N_k \leqslant i_k s^{1-k/n} N_n^{k/n} \qquad (k = 1, 2, \ldots, n-1). \quad (2.7.2)$$

It has been shown by Mardzhanishvili[19] that the following conditions must be fulfilled for system (2.7.1) to be solvable:

$$\left|\begin{matrix} N_n(n-1)^n \cdots 1^n \\ N_{n-1}(n-1)^{n-1} \cdots 1^{n-1} \\ \cdots\cdots\cdots\cdots \\ N_1\, n-1 \cdots 1 \end{matrix}\right| \equiv 0 \pmod R; \ldots;$$

$$\left|\begin{matrix} n^n(n-1)^n \cdots N_n \\ n^{n-1}(n-1)^{n-1} \cdots N_{n-1} \\ \cdots\cdots\cdots\cdots \\ n \quad n-1 \cdots N_1 \end{matrix}\right| \equiv 0 \pmod R, \qquad (2.7.3)$$

where

$$R = \left|\begin{matrix} n^n(n-1)^n \cdots 1^n \\ n^{n-1}(n-1)^{n-1} \cdots 1^{n-1} \\ n \quad n-1 \cdots 1 \end{matrix}\right|.$$

Let us prove this assertion.

Suppose that equations (2.7.1) hold. We denote by D_n the first of the determinants appearing in the left-hand side of (2.7.3).

On applying elementary rules for evaluating determinants, we find successively:

$$D_n = \begin{vmatrix} \sum x_\nu^n (n-1)^n \cdots 1^n \\ \cdots\cdots\cdots\cdots \\ \sum x_\nu (n-1) \cdots 1 \end{vmatrix}$$

$$= (n-1)! \begin{vmatrix} \sum x_\nu^n (n-1)^{n-1} \cdots 1 \\ \cdots\cdots\cdots\cdots \\ \sum x_\nu \quad 1 \cdots 1 \end{vmatrix} = (n-1)!$$

$$\times (n-2)! \begin{vmatrix} \sum x_\nu^{n-1}(x_\nu - 1)(n-1)^{n-2} \cdots 2^{n-2} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \sum x_\nu (x_\nu - 1) \quad 1 \cdots 1 \end{vmatrix} = (n-1)!$$

$$\times (n-2)!(n-3)! \begin{vmatrix} \sum x_\nu^{n-2}(x_\nu - 1)(x_\nu - 2)(n-1)^{n-3} \cdots 3^{n-3} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \sum x_\nu (x_\nu - 1)(x_\nu - 2) \quad 1 \cdots 1 \end{vmatrix}$$

$$= \cdots = (n-1)!(n-2)! \ldots 2!\,1! \sum x_\nu (x_\nu - 1) \cdots (x_\nu - n + 1)$$

$$\equiv 0 \pmod{R}$$

We shall prove the following theorem: given a sufficiently large $S = S(n)$ and a set of numbers $N_1, N_2, \ldots, N_n$, satisfying conditions (2.7.2) and (2.7.3), the system of equations (2.7.1) is solvable. An elementary solution of this problem was obtained in 1948 by Emel'yanov[8] in analogy with the solution of Waring's problem described above. We shall follow Emel'yanov's method here.

8. Sequences of Integral-valued Vectors and their Densities

With a view to solving the Hilbert–Kamke problem by elementary means we propose to carry over to integral-valued vectors the ideas of density and addition of sequences of integers described above.

We shall utilize the following notation in future: $\bar{m}$, $\bar{A}, \bar{B}, \ldots$, $\bar{T}, \bar{Z}$ are vectors with non-negative integral-valued components in space E^n;

$\{\bar{A}\}$, $\{\bar{T}\}$, $\{\bar{m}\}$ are sets of integral-valued vectors of E^n;

$\{\bar{N}\}$ is the system of all integral-valued vectors of E^n;

$\{\overline{M}\}$ is the system of integral-valued vectors of E^n whose components $M_1, M_2, \ldots, M_n$ satisfy conditions (2.7.2) and (2.7.3).

If the vector $\bar{a}$ belongs to the set $\{\bar{A}\}$, we shall write as usual: $\bar{a} \in \{\bar{A}\}$. If all the components of two vectors $\bar{a}$ and $\bar{b}$, i.e. a_k and b_k $(k = 1, 2, \ldots, n)$, are connected by the relationships $a_k \leqslant b_k$ $(k = 1, 2, \ldots, n)$, we shall write: $\bar{a} \subseteq \bar{b}$ ($\bar{a}$ "does not exceed" $\bar{b}$). The addition of sequences of integral-valued vectors is defined like the addition of numerical sequences in Schnirelmann's definition:

$\{\bar{A}\} + \{\bar{B}\} = \{\bar{c}\}$, where $\{\bar{c}\}$ is the set of vectors $\bar{c}$ of the form $\bar{c} = \varepsilon_1\bar{a} + \varepsilon_2\bar{b}\,;\, \bar{a} \in \{\bar{A}\}\,;\, \bar{b} \in \{\bar{B}\}\,;$

$$\varepsilon_1 = 0;\ 1;\ \varepsilon_2 = 0;\ 1;\ \varepsilon_1 + \varepsilon_2 \neq 0\,.$$

Addition of vectors is understood in the usual sense. The congruence of two vectors modulo a scalar: $\bar{a} \equiv \bar{b} \pmod{q}$, implies: $a_k \equiv b_k \pmod{q}$ $(k = 1, 2, \ldots, n)$.

If $\bar{a} = \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_1 \end{pmatrix}$, then $(+)\,\bar{a}$ denotes the vector $\begin{pmatrix} |a_n| \\ |a_{n-1}| \\ \vdots \\ |a_1| \end{pmatrix}$.

The definition of the density of a sequence of integral-valued vectors of E^n is similar to Schnirelmann's definition for sequences of integers:

$$D\{\bar{A}\} = \inf \frac{r_{\bar{m}}\{\bar{A}\}}{r_{\bar{m}}\{\bar{N}\}}\,,$$

where $\bar{m}$ runs over all $\{\bar{N}\}$, $r_m\{\bar{A}\}$ is the number of vectors of $\{\bar{A}\}$ not exceeding $\bar{m}$, $r_{\bar{m}}\{\bar{N}\}$ is the number of vectors of $\{\bar{N}\}$ not exceeding $\bar{m}$; $\{\bar{A}\}$ is called the basis of $\{\bar{N}\}$ of order l, if each vector $\bar{m} \in \{\bar{N}\}$ can be written in the form

$$\bar{m} = \bar{a}_1 + \bar{a}_2 + \cdots + \bar{a}_k,$$

where $\bar{a}_i \in \{\bar{A}\}$ $(i = 1, 2, \ldots, k;\ k \leqslant l)$.

9. Some Lemmas

We shall here prove some lemmas, very similar to the lemmas of §§ 3 and 4.

LEMMA 1. *Let* $F_1 = \{\bar{P}_i\}$, $F_2 = \{\bar{Q}_j\}$ *be two finite systems of integral-valued vectors of* E^n, *where each* $\bar{P}_j \in F_1$ *is repeated* $a_j \geqslant 1$ *times, and each vector* $\bar{Q}_j \in F_2$ *is repeated* b_j *times. Then the N.S.E.*

$$P_i + Q_j = Z, \tag{2.9.1}$$

where $\bar{Z}$ is any fixed integral-valued vector, does not exceed half the sum of the N.S.E.

$$\bar{P}_i - \bar{P}_j = \bar{O} \quad \text{and} \quad \bar{Q}_i - \bar{Q}_j = \bar{O}, \tag{2.9.2}$$

where $\bar{O} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ *is the zero vector, which we shall in future denote simply by "O".*

For, equation (2.9.1) associates each i with a definite j, i.e. each of the a_i equal vectors $\bar{P}_i$ with any of the b_j equal $\bar{Q}_j$, which gives

$$a_i b_j \leqslant \frac{a_i^2 + b_j^2}{2},$$

while, since $\bar{P}_{i_1} \neq \bar{P}_{i_2}$ when $i_1 \neq i_2$, the vectors $\bar{Q}_{j_1}$ and $\bar{Q}_{j_2}$ corresponding to them by virtue of equation (2.9.1) are likewise not equal, so that the N.S.E. (2.9.1) $\leqslant \frac{1}{2}(\Sigma a_i^2 + \Sigma b_j^2)$, where $\Sigma a_i^2 + \Sigma b_j^2$ is the sum of the N.S.E. (2.9.2).

LEMMA 2. *Given $2^s k_0$ finite systems of vectors $G_i = \{\bar{M}_{ij}\}$ $(i = 1, 2, \ldots, 2^s k_0)$, each vector $\bar{M}_{ij} \in G_i$ being repeated $a_{ij} \geqslant 1$ times.*

We consider the N.S.E.

$$\left.\begin{array}{c} \bar{M}_1 + \bar{M}_2 + \cdots + \bar{M}_{2^s k_0} = \bar{Z}; \\ \bar{M}_i \in G_i \qquad (i = 1, 2, \ldots 2^s k_0. \end{array}\right\} \tag{2.9.3}$$

We introduce the 2^s systems of integral-valued vectors $\{\bar{Y}\}$ of the form

$$\bar{Y}_j = \bar{M}_{(j-1)k_0+1} + \bar{M}_{(j-1)k_0+2} + \cdots + \bar{M}_{(j-1)k_0+k_0} \qquad (j = 1, 2, \ldots, 2^s), \tag{2.9.4}$$

Then the N.S.E. (2.9.3) *does not exceed the sum of the numbers of solutions of the independent equations*

$$\left.\begin{array}{l} \bar{Y}_1^{(1)} + \bar{Y}_1^{(2)} + \cdots + \bar{Y}_1^{(2^{s-1})} - \bar{Y}_1^{(2^{s-1}+1)} - \cdots - \bar{Y}_1^{(2^s)} = 0, \\ \bar{Y}_2^{(1)} + \bar{Y}_2^{(2)} + \cdots + \bar{Y}_2^{(2^{s-1})} - \bar{Y}_2^{(2^{s-1}+1)} - \cdots - \bar{Y}_2^{(2^s)} = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \bar{Y}_{2^s}^{(1)} + \bar{Y}_{2^s}^{(2)} + \cdots + \bar{Y}_{2^s}^{(2^{s-1})} - \bar{Y}_{2^s}^{(2^{s-1}+1)} - \cdots - Y_{2^s}^{(2^s)} = 0, \end{array}\right\} \tag{2.9.5}$$

where all the variables of the jth row run independently over the values (2.9.4) *with the relevant number of repetitions.*

In fact, on considering the two systems of vectors

$$\{\overline{M}_1 + \overline{M}_2 + \cdots \overline{M}_{2^{s-1}k_0}\} \quad \text{and} \quad \{\overline{M}_{2^{s-1}k_0+1} + \cdots + \overline{M}_{2^s k_0}\},$$

we can apply to (2.9.3) Lemma 1 and obtain the two equations

$$(\overline{M}_1 - \overline{M}'_1) + (\overline{M}_2 - \overline{M}'_2) + \cdots + (\overline{M}_{2^{s-1}k_0} - \overline{M}'_{2^{s-1}k_0}) = 0,$$

$$(\overline{M}_{2^{s-1}k_0+1} - \overline{M}'_{2^{s-1}k_0+1}) + \cdots + (\overline{M}_{2^s k_0} - \overline{M}'_{2^s k_0}) = 0.$$

We again apply Lemma 1 to each of these equations, and in this way, after s steps, arrive at the assertion of Lemma 2.

LEMMA 3. *The N.S.E.*

$$h_1\bar{z}_1 + h_2\bar{z}_2 + \cdots + h_\nu\bar{z}_\nu = \overline{m}, \tag{2.9.6}$$

where $(+)\ \bar{z}_i \subseteq \overline{T}$, $|h_i| \leqslant p$, p *is a fixed number,*

$$\overline{T} = \begin{pmatrix} T_n \\ T_{n-1} \\ \vdots \\ T_1 \end{pmatrix}, \quad T_i \geqslant 0, \quad p \leqslant T_i\,(i = 1, 2, \ldots, n),$$

the h_i *are fixed,* $(h_1, h_2, \ldots, h_\nu) = 1$, *and they do not exceed*

$$B_\nu \frac{(r_{\overline{T}}\{\overline{N}\})^{\nu-1}}{\max |h_i|^n}. \tag{2.9.7}$$

We have for $\nu = 2$: $h_1\bar{z}_1 + h_2\bar{z}_2 = \overline{m}$, $|h_1| > |h_2|$, $(h_1, h_2) = 1$ and $h_2\bar{z}_2 = \overline{m} \pmod{h_1}$. The vector $\bar{z}_2$ is defined uniquely modulo h_1 and hence can take not more than $Br_{\overline{T}}\{\overline{N}\}/|h_1|^n$ values. To each vector $\bar{z}_2$ there corresponds not than one $\bar{z}_1$.

Now let the lemma be true for $\nu - 1$ and let $\max |h_i| = |h_\nu|$ $(i = 1, 2, \ldots, \nu)$. Let $(h_1, h_2, \ldots, h_{\nu-1}) = \delta$ and

$$\max\left\{\frac{|h_1|}{\delta}, \frac{|h_2|}{\delta}, \ldots, \frac{|h_{\nu-1}|}{\delta}\right\} = H'.$$

Then the N.S.E. $h_1\bar{z}_1 + h_2\bar{z}_2 + \ldots + h_{\nu-1}\bar{z}_{\nu-1} = \delta\overline{Y}$ does not exceed $B(r_{\overline{T}}\{\overline{N}\})^{\nu-2}/(H')^n$, where each value $(+)\ \overline{Y} \subseteq \nu H'\overline{T}$. Further, obviously, $(\delta, h_\nu) = 1$ and the N.S.E.

$$\delta\overline{Y} + h_\nu\bar{z}_\nu = \overline{m}$$

will be

$$B\frac{(r_{\overline{T}}\{\overline{N}\})^{\nu-2}}{(H')^n}\frac{r_{\nu H'\overline{T}}\{\overline{N}\}}{|h_\nu|^n} = B\frac{(r_{\overline{T}}\{\overline{N}\})^{\nu-1}}{|h_\nu|^n},$$

which is what we wanted to prove.

LEMMA 4. *The sum of the N.S.E.*

$$h_1\bar{z}_1 + h_2\bar{z}_2 + \cdots + h_\nu\bar{z}_\nu = 0, \tag{2.9.8}$$

where $(+)\ \bar{z}_i \subseteq \bar{T}$; h_i *run independently over the interval* $[-2p, 2p]$, *the components of the vector* $\bar{T}$, i.e. $T_k \leqslant Bp^k$, $T_k \geqslant 2p$, *will be not greater than*

$$B\frac{(r_{\bar{T}}\{\bar{N}\})^{\nu-1}}{p^n}p^\nu \quad \text{for} \quad \nu = 4n^2.$$

For systems $\{h_1, h_2, \ldots, h_\nu\}$, containing not more than two $h_i \neq 0$, we take the trivial upper bound of the N.S.E. (2.9.8): $B(r_{\bar{T}}\{\bar{N}\})^{\nu-1}$, so that the sum of the numbers of solutions is in this case not greater than

$$B(r_{\bar{T}}\{\bar{N}\})^{\nu-1}p^2 = \frac{B(r_{\bar{T}}\{\bar{N}\})^{\nu-1}}{p^n}p^\nu\frac{p^n}{p^{\nu-2}} = B\frac{(r_{\bar{T}}\{\bar{N}\})^{\nu-1}}{p^n}p^\nu. \tag{2.9.9}$$

We classify the remaining systems $\{h_1, h_2, \ldots, h_\nu\}$ according to their greatest common divisor δ. In the case of those systems for which $(h_1, h_2, \ldots, h_\nu) = 1$, we split the interval $[1, 2p]$ into subintervals $[2p/2^{m+1}, 2p/2^m]$, the number of which is not greater than $B \ln p$.

Let $H = \max|h_i|$. We consider all the systems for which $H \in [2p/2^{m+1}, 2p/2^m]$. The sum of the N.S.E. (2.9.8) for all such equations does not exceed, by Lemma 3,

$$B\frac{(r_{\bar{T}}\{\bar{N}\})^{\nu-1}}{H^n}\left(3\cdot\frac{2p}{2^m}\right)\left(\frac{2p}{2^m}\right)^{\nu-1} \leqslant B\frac{(r_{\bar{T}}\{\bar{N}\})^{\nu-1}}{p^n}2^{mn}p^\nu 2^{-m\nu}$$
$$= B_\nu\frac{(r_{\bar{T}}\{\bar{N}\})^{\nu-1}}{p^n}p^\nu\frac{1}{2^{m(\nu-n)}}. \tag{2.9.10}$$

On summing (2.9.10) over all $m > 0$, we find in this case that the sum of the N.S.E. (2.9.8) does not exceed

$$B\frac{(r_{\bar{T}}\{\bar{N}\})^{\nu-1}}{p^n}p^\nu. \tag{2.9.11}$$

On now considering the systems with $(h_1, h_2, \ldots, h_\nu) = \delta$, it can be seen that the expression for the sum of the numbers of solutions of (2.9.8) is obtained in this case by replacing p in (2.9.11) by p/δ, as a result of which we obtain for fixed δ: the sum of the N.S.E. (2.9.8) does not exceed

$$B\frac{(r_{\bar{T}}\{\bar{N}\})^{\nu-1}}{p^n}\frac{p^\nu}{\delta^{\nu-n}}. \tag{2.9.12}$$

Since $\nu > n + 2$, on summation over all values of δ we obtain the convergent series $\sum_{\delta} 1/\delta^{\nu-n}$, i.e. a constant, which proves the assertion of the lemma for every case.

We now prove a fundamental lemma.

LEMMA 5. *Let the vector* $\bar{m} \in \{\bar{N}\}$, $\bar{m} \subseteq \bar{T}$, $\bar{T} \in \{\bar{N}\}$. *We consider the set of vectors*

$$\bar{T} = \begin{pmatrix} T_n \\ T_{n-1} \\ \vdots \\ T_1 \end{pmatrix}; \quad \{\overline{F(x)}\} = \left\{ \begin{pmatrix} f_n(x) \\ f_{n-1}(x) \\ \vdots \\ f_1(x) \end{pmatrix} \right\}, \tag{2.9.13}$$

where $f_j(x) = a_{0j}x^j + a_{1j}x^{j-1} + \ldots + a_{jj}$ *is a polynomial with integral coefficients,* $a_{0j} \neq 0$ *and depends only on* j; x *runs over all the integers*

$$a_{ij} = a_{ij}(j, T_j)\ (i > 1), \quad |a_{ij}| \leqslant B_n T_n^{1/n} = B_n p. \tag{2.9.14}$$

Then there exists $K = K(n)$ *such that the N.S.E.*

$$\overline{F(x_1)} + \overline{F(x_2)} + \cdots + \overline{F(x_k)} = \bar{m}, \tag{2.9.15}$$

where $|x_i| \leqslant p$ *does not exceed*

$$B_n \frac{p^K}{p^{n^2}}. \tag{2.9.16}$$

10. Proof of the Fundamental Lemma

For $n = 1$, i.e. for vectors of space E^1, which are ordinary linear polynomials $ax + b$, the lemma is obvious, and $K(1) = 1$.

Suppose that the lemma has been proved for all vector-polynomials of space E^{n-1}, and let us denote the corresponding $K(n-1) = K'$. We now put $K = 2^{(4\log_2 K')}\, 4n$ and consider the equation

$$\overline{F(x_1)} + \overline{F(x_2)} + \cdots + \overline{F(x_K)} = \bar{m}, \tag{2.10.1}$$

$$|x_i| \leqslant p = T_n^{1/n} \qquad (i = 1, 2, \ldots, K); \quad m \in \{\bar{N}\}.$$

On applying Lemma 1 to (2.10.1), we can replace this equation by

$$\{\overline{F(x_1)} - \overline{F(y_1)}\} + \{\overline{F(x_2)} - \overline{F(y_2)}\} + \cdots + \{\overline{F(x_{K/2})} - \overline{F(y_{K/2})}\} = 0, \tag{2.10.2}$$

where x_i, y_i are independent and lie in the interval $[-p, p]$. We put $x_i = y_i + h_i$, where y_i and h_i run independently over the

interval $[-2p, 2p]$, as a result of which the number of solutions of equation (2.10.2) can only increase. For each fixed system of values $(h_1, h_2, \ldots, h_{K/2})$ we obtain the equation

$$h_1\overline{\varphi_1(y_1)} + h_2\overline{\varphi_2(y_2)} + \cdots + h_{K/2}\overline{\varphi_{K/2}(y_{K/2})} = 0, \qquad (2.10.3)$$

where $\overline{\varphi_i(y_i)}$ is a vector-polynomial with components

$$(na_{0n}y_i^{n-1} + l_{1n}^{(i)}y_i^{n-2} + \cdots + l_{n-1n}^{(i)}; \cdots; a_{01},$$

where $|l_{\gamma j}^{(i)}| \leqslant 2^\gamma p^\gamma$ $(\gamma = 1, 2, \ldots, n-1)$. These vector-polynomials $\overline{\varphi_i(y_i)}$ all have a constant first component a_{01}, and hence, on considering all the vectors $\overline{\varphi_i(y_i)}$ lying in E^{n-1}, we obtain the supplementary condition

$$h_1 + h_2 + \cdots + h_{K/2} = 0. \qquad (2.10.4)$$

We shall now consider the system of vector-polynomials $h_i\overline{\varphi_i(y_i)}$ as a system G_i of Lemma 2, where $K_0 = 4n$,

$$2^s = 2^{[4\log_2 K']-1}. \qquad (2.10.5)$$

This gives us 2^s equations of the form

$$\begin{aligned} h_1\{\overline{\varphi_1(y_1^{(1)})} + \overline{\varphi_1(y_1^{(2)})} + \cdots + \overline{\varphi_1(y_1^{(2^{s-1})})} - \overline{\varphi_1(y_1^{(2^{s-1}+1)})} \\ - \cdots - \overline{\varphi_1(y_1^{(2^s)})}\} + \cdots + h_{4n}\{\overline{\varphi_{4n}(y_{4n}^{(1)})} \\ + \cdots - \overline{\varphi_{4n}(y_{4n}^{(2^s)})}\} = 0, \end{aligned} \qquad (2.10.6)$$

where the $y_i^{(j)}$ run independently over the interval $[-2p, 2p]$.

The expressions contained in each of the $4n$ curly brackets are vectors $\bar{z}_i$ of E^{n-1}, satisfying the condition

$$(+)\,\bar{z} \subseteq \begin{pmatrix} c_{n-1}p^{n-1} \\ \vdots \\ c_1 p \end{pmatrix},$$

and, in accordance with our inductive assumption, each such vector $\bar{z}_i$ is repeated not more than $Bp^{2s}/p^{(n-1)^2}$ times.

We thus arrive at 2^s linear equations of the form

$$\left.\begin{aligned} &h_1\bar{z}_1 + h_2\bar{z}_2 + \cdots + h_{4n^2}\bar{z}_{4n} = 0, \\ &h_{4n+1}\bar{z}_{4n+1} + h_{4n+2}\bar{z}_{4n+2} + \cdots + h_{4n+4n}\bar{z}_{4n+4n} = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &h_{K/2-4n+1}\bar{z}_{K/2-4n+1} + \cdots + h_{K/2}\bar{z}_{K/2} = 0, \end{aligned}\right\} \qquad (2.10.7)$$

where

$$(+)\,\bar{z} \subseteq \begin{pmatrix} c_{n-1}p^{n-1} \\ \vdots \\ c_1 p \end{pmatrix}$$

and each vector $\bar{z}_i$ is repeated not more than $Bp^{2s}/p^{(n-1)^2}$ times; the numbers h_i meantime satisfy the condition

$$\sum_{i=1}^{K/2} h_i = 0. \tag{2.10.8}$$

If we oblige h_i to run over all values, including $(0, 0, \ldots, 0)$ and dispense with condition (2.10.8), we shall merely increase the sum of the N.S.E. (2.10.7). But if we now apply Lemma 4, we find that: the sum of the N.S.E. (2.10.7), where no repetitions of a vector are allowed, will not exceed the number

$$B\frac{(r_{\bar{T}}\{\bar{N}\})^{4n-1}}{p^n}(p^{4n})^{2^s}. \tag{2.10.9}$$

Since each $\bar{z}_i$ is repeated not more than $Bp^{2s}/p^{(n-1)^2}$ times, the required sum of the N.S.E. (2.10.7) when repetitions are allowed will not exceed

$$\frac{B(r_{\bar{T}}\{\bar{N}\})^{4n-1}}{p^n}(p^{4n})^{2^s}\left(\frac{p^{2^s}}{p^{(n-1)^2}}\right)^{4n}$$

$$\leqslant B_n \frac{p^{\frac{n(n+1)}{2}(4n-1)}p^{4n2^s}p^{4n2^s}}{p^{n^2}p^{4n(n-1)^2}} \leqslant B_n \frac{p^K}{p^{n^2}}, \tag{2.10.10}$$

which is what we wished to prove.

We now consider vector-polynomials of the form

$$\overline{F_1(x_i)} = \begin{pmatrix} x_i^n \\ x_i^{n-1} \\ \vdots \\ x_i \end{pmatrix} \tag{2.10.11}$$

and the aggregate of them $\{\overline{F_1(x_i)}\}$, where x_i runs over all the positive integers. Obviously, these vectors satisfy all the conditions of the fundamental lemma. Let us show that the density of $\{\overline{F_1(x_i)}\}$ is positive.

If we fix the vector $\bar{m} \in \{\bar{N}\}$, and consider all integral-valued vectors of E^n, including $\overline{F_1(x_i)} \subseteq \bar{m}$, in accordance with the

fundamental lemma we can find a number $K(n)$. If we further consider all the $\overline{F_1(x_i)} \subseteq \bar{m}$ with numbers of repetitions a_j, satisfying the condition

$$\sum_{j=1}^{r_{\bar{m}}\{\overline{F_1(x_i)}\}} a_j = B_n m_n^{K/n}, \tag{2.10.12}$$

we obtain

$$\sum_{j=1}^{r_{\bar{m}}\{\overline{F_1(x_i)}\}} a_j \leqslant B_n \frac{m_n^{K/n}}{m_n^n} r_{\bar{m}}\{\overline{F_1(x_i)}\},$$

whence

$$r_{\bar{m}}\{F_1(x_i)\} \geqslant c \frac{\sum_{j=1}^{r_{\bar{m}}\{\overline{F_1(x_i)}\}} a_i m_n^n}{m_n^{K/n}} \geqslant c_n m_n^n;$$

$$\frac{r_{\bar{m}}\{\overline{F_1(x_i)}\}}{r_{\bar{m}}\{\bar{N}\}} \geqslant c_n' \frac{m_n^n}{r_{\bar{m}}\{\bar{N}\}} \geqslant c_n'' > 0, \tag{2.10.13}$$

which is what we wished to prove.

In additive problems of the type of Waring's problem, a well known advantage of the Schnirelmann concept of density of a numerical sequence is that we are freed from the necessity of investigating the arithmetical structure of the sequence, provided the density be positive. No such situation arises in the case of integral-valued vectors of E^n, $n > 1$. For, let $\bar{A}$ be a subset of $\{\bar{N}\}$. We denote by $\mathfrak{A}$ the arithmetical characteristic of $\{\bar{A}\}$, i.e. the rule which enables us to decide, in a finite number of operations, whether a previously assigned vector $\bar{m}$ belongs to $\{\bar{A}\}$. Let $\{\bar{B}\}$ be a set of vectors, all of which are different, and are of the form

$$\bar{B} = \sum_{i=1}^{K} \bar{A}_i, \ \bar{A}_i \in \{\bar{A}\},$$

where K runs over all the positive integers. The set $\{\bar{B}\}$ now has a non-trivial arithmetical characteristic $\mathfrak{B}$. In fact, suppose say that the characteristic $\mathfrak{A}$ is such that it excludes from $\{\bar{A}\}$ all vectors of the form

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ m \end{pmatrix};$$

then it is obvious that all these vectors will likewise not belong to $\{\bar{B}\}$, i.e. the set $\{\bar{A}\}$, provided it also has positive density (which is evidently possible, since it can contain, for example, all the vectors other than those indicated above), is not a basis for $\{\bar{N}\}$. Consequently it is only possible to pose the question of determining the property of the set $\mathfrak{B}$ if the system of vectors $\{\bar{A}\}$ with the property $\mathfrak{A}$ is given, and next, whether or not $\{\bar{A}\}$ is a basis of finite order for $\{\bar{B}\}$, if $D\{\bar{A}\} > 0$. We shall only consider vectors

$$\bar{m} = \begin{pmatrix} m_n \\ m_{n-1} \\ \vdots \\ m_1 \end{pmatrix},$$

satisfying the condition

$$m_1 \leqslant m_2 \leqslant \cdots \leqslant m_n. \tag{2.10.14}$$

We define the density of the set $\{\bar{A}\}$ with respect to $\{\bar{B}\}$ as

$$D_{\{\bar{B}\}}\{\bar{A}\} = \inf \frac{r_{\bar{m}}\{\bar{A}\}}{r_{\bar{m}}\{\bar{B}\}}, \tag{2.10.15}$$

where $\bar{m}$ runs over all integral-valued vectors $\{\bar{N}\}$, and prove the lemma

$$D_{\{\bar{B}\}}(\{\bar{A}\} + \{\bar{A}\}) \geqslant 2D_{\{\bar{B}\}}\{\bar{A}\} - D^2_{\{\bar{B}\}}\{\bar{A}\}. \tag{2.10.16}$$

Let $\bar{m} \in \{\bar{B}\}$, and let us order $\{\bar{A}\}$ and $\{\bar{B}\}$ so that any given vector never "exceeds" the succeeding vector. Now, if $\{\bar{B}\} \neq \{\bar{A}\}$, with a given $\bar{m} \in \{\bar{B}\}$ we can find an $\bar{a}_1 \in \{\bar{A}\}$ and $\bar{a}_2 \in \{\bar{A}\}$ such that $\bar{a}_1 \subset \bar{m}$, $\bar{a}_2 \subseteq \bar{m}$ and there is at any rate one vector $\bar{b} \in \{\bar{B}\}$ such that $\bar{a}_1 \subset \bar{b} \subset \bar{a}_2$. The number of all such vectors $\bar{b} \subseteq \bar{m}$ will be $r_{\bar{m}}\{\bar{B}\} - r_{\bar{m}}\{\bar{A}\}$. Let $D_{\{\bar{B}\}}\{\bar{A}\} = \alpha > 0$; then $r_{\bar{a}_2 - \bar{a}_1}\{\bar{A}\} > \alpha r_{\bar{a}_2 - \bar{a}_1}\{\bar{B}\}$, and consequently the interval between vectors $\bar{a}_1$ and $\bar{a}_2$ contains not less than $\alpha r_{\bar{a}_2 - \bar{a}_1}\{\bar{B}\}$ vectors

$$\bar{c} \in \{\bar{A}\} + \{\bar{A}\}.$$

If we now sum them over all such intervals, we get

$$\begin{aligned} r_{\bar{m}}(\{\bar{A}\} + \{\bar{A}\}) &\geqslant r_{\bar{m}}\{\bar{A}\} + \alpha \sum r_{\bar{a}_i - \bar{a}_{i-1}}\{\bar{B}\} \\ &= r_{\bar{m}}\{\bar{A}\} + \alpha(r_{\bar{m}}\{\bar{B}\} - r_{\bar{m}}\{\bar{A}\}), \end{aligned}$$

or

$$r_{\bar{m}}(\{\bar{A}\} + \{\bar{A}\}) \geqslant 2\alpha r_{\bar{m}}\{\bar{B}\} - \alpha r_{\bar{m}}\{\bar{A}\}.$$

Hence

$$\frac{r_{\bar{m}}(\{\bar{A}\} + \{\bar{A}\})}{r_{\bar{m}}\{\bar{B}\}} \geqslant 2\alpha - \frac{\alpha r_{\bar{m}}\{\bar{A}\}}{r_{\bar{m}}\{\bar{B}\}},$$

or

$$D_{\{\bar{B}\}}(\{\bar{A}\} + \{\bar{A}\}) \geqslant 2\alpha - \alpha^2.$$

Further, it is obvious that, if we take a sufficiently large number S, we obtain $D_{\{\bar{B}\}}\left(\sum_{i=1}^{S}\{\bar{A}\}\right) > \frac{1}{2}$, whence it follows at once that $\{\bar{A}\}$ is a basis of order $2S$ for the set of vectors $\{\bar{B}\}$.

Let us apply this lemma to the set of vectors $\{\overline{F_1(x_i)}\}$. This set satisfies the conditions of the lemma and its density is positive. Further, it is well known that every vector

$$\begin{pmatrix} N_n \\ N_{n-1} \\ \vdots \\ N_1 \end{pmatrix},$$

representable as a sum of vectors of $\{\overline{F_1(x_i)}\}$, must satisfy the necessary condition (2.7.3). This condition is in fact the property $\mathfrak{B}$ of the set of all vectors

$$\begin{pmatrix} N_n \\ N_{n-1} \\ \vdots \\ N^1 \end{pmatrix},$$

representable as a sum of a certain number of terms of $\{\overline{F_1(x_i)}\}$. Since $D_{\{\bar{B}\}}\{\overline{F_1(x_i)}\} \geqslant D\{\overline{F_1(x_i)}\} > 0$, the fundamental theorem is thereby proved.

CHAPTER 3

THE PROBLEM OF THE DISTRIBUTION OF PRIMES

1. Numerical Functions and the Relationships among them. Estimation of the Number of Primes in a Segment of the Series of Positive Integers

Elementary methods dealing with the distribution of primes, based on investigation of the properties of the Mangoldt and Möbius numerical functions $\Lambda(n)$ and $\mu(n)$:

$$\Lambda(n) = \begin{cases} \ln p, & n = p^k \\ 0, & n \neq p^k \end{cases}, \quad \mu(n) = \begin{cases} +1, & n = 1, \\ (-1)^k, & n = p_1 \cdots p_k, \\ 0, & p^2 / n, \end{cases}$$

where p and p_k are primes, have recently undergone substantial development as a result of the work of Atle Selberg and his followers. We shall derive a number of well known properties of these functions which we shall require later.

First of all,

$$\sum_{d|n} \mu(d) = \prod_{p|n} [1 + \mu(p)] = \begin{cases} 1, & n = 1, \\ 0, & n > 1, \end{cases} \tag{3.1.1}$$

and in the case of arbitrary $U(x)$, if

$$V(n) = \sum_{d|n} U(d),$$

then

$$\sum_{k|n} \mu(k) V\left(\frac{n}{k}\right) = \sum_{k|n} \mu(k) \sum_{d k|n} U(d)$$
$$= \sum_{d|n} U(d) \sum_{k|n/d} \mu(k) = U(n). \tag{3.1.2}$$

It follows from this last relationship, since

$$\ln n = \sum_{d|n} \Lambda(d),$$

that

$$\Lambda(n) = \sum_{d|n} \mu(d) \ln \frac{n}{d} = -\sum_{d|n} \mu(d) \ln d.$$

Further, if

$$\varrho(n) = \begin{cases} 1, & n \geqslant 1, \\ 0, & n < 1, \end{cases}$$

then

$$[x] = \sum_{n=1}^{\infty} \varrho\left(\frac{x}{n}\right), \quad \varrho(x) = \sum_{n \leqslant x} \sum_{d|n} \mu(d)$$

$$= \sum_{d \leqslant x} \mu(d) \left[\frac{x}{d}\right]. \tag{3.1.3}$$

Hence it follows that, when $x \geqslant 1$,

$$1 = x \sum_{n \leqslant x} \frac{\mu(n)}{n} - \sum_{n \leqslant x} \mu(n) \left\{\frac{x}{n}\right\}$$

or that

$$\left| \sum_{n \leqslant x} \frac{\mu(n)}{n} \right| \leqslant 1. \tag{3.1.4}$$

Similarly,

$$\psi(x) = \sum_{n \leqslant x} \Lambda(n) = \sum_{n \leqslant x} \sum_{d|n} \mu(d) \ln \frac{x}{d} - \ln x$$

$$= \sum_{n \leqslant x} \left[\frac{x}{n}\right] \mu(n) \ln \frac{x}{n}$$

$$= -\sum_{n \leqslant x} \left\{\frac{x}{n}\right\} \ln \frac{x}{n} \mu(n) + x \sum_{n \leqslant x} \frac{\mu(n)}{n} \ln \frac{x}{n}.$$

But, when $q > 0$,

$$\sum_{n \leqslant x} \ln^q \frac{x}{n} = \int_1^x \ln^q \frac{x}{t}\, dt + O(\ln^q x)$$

$$= x \int_1^x \ln^q t \frac{dt}{t^2} + O(\ln^q x) = O(x), \tag{3.1.5}$$

so that

$$\left| \sum_{n \leqslant x} \left\{\frac{x}{n}\right\} \mu(n) \ln \frac{x}{n} \right| = O(x).$$

Further, since

$$\sum_{k\leqslant t}\frac{1}{k}=\ln t+c+O\left(\frac{1}{t}\right),$$

we have

$$\sum_{n\leqslant x}\frac{\mu(n)}{n}\ln\frac{x}{n}=\sum_{n\leqslant x}\frac{\mu(n)}{n}\left[\sum_{k\leqslant x/n}\frac{1}{k}-c+O\left(\frac{n}{x}\right)\right]$$

$$=\sum_{n\leqslant x}\frac{1}{n}\sum_{d\mid n}\mu(d)-c\sum_{n\leqslant x}\frac{\mu(n)}{n}+O(1)=O(1).$$

Hence

$$\psi(x)=\sum_{n\leqslant x}\Lambda(n)=O(x). \tag{3.1.6}$$

On using Stirling's formula, we have

$$\sum_{n\leqslant x}\ln n=x\ln x-x+O(\ln x)=\sum_{n\leqslant x}\sum_{d\mid n}\Lambda(d)$$

$$=\sum_{d\leqslant x}\Lambda(d)\left[\frac{x}{d}\right]=x\sum_{n\leqslant x}\frac{\Lambda(n)}{n}-\sum_{d\leqslant x}\Lambda(d)\left\{\frac{x}{d}\right\}$$

$$=x\sum_{n\leqslant x}\frac{\Lambda(n)}{n}+O(x)$$

or alternatively,

$$\sum_{n\leqslant x}\frac{\Lambda(n)}{n}=\ln x+O(1). \tag{3.1.7}$$

This last formula not only shows that the number of primes is infinite, but also gives a representation of the number of them in the interval $[1, x]$. For,

$$\sum_{n\leqslant x}\frac{\Lambda(n)}{n}=\sum_{p\leqslant x}\frac{\ln p}{p}+\sum_{p^2\leqslant x}\frac{\ln p}{p^2}+\cdots$$

$$=\sum_{p\leqslant x}\frac{\ln p}{p}+O(1)=\ln x+O(1),$$

in other words, not only is the number of primes infinite, but the series with terms $(\ln p)/p$ is divergent. A more accurate estimate of the number of primes can be obtained by considering the lower bound of $\psi(x)$. We have further:

$$x\ln x-x+O(\ln x)=\sum_{n\leqslant x}\Lambda(n)\left[\frac{x}{n}\right]$$

$$=\sum_{d\leqslant x}\Lambda(d)\sum_{k\leqslant x/d}1=\sum_{d\leqslant x}\sum_{k\leqslant x/d}\Lambda(k)=\sum_{1}^{\infty}\psi\left(\frac{x}{n}\right),$$

if use is made of the transformation

$$\sum_{n\leqslant x} U(n) \sum_{m\leqslant x/n} V(m) = \sum_{n\leqslant x} V(n) \sum_{m\leqslant x/n} U(m). \tag{3.1.8}$$

Hence, following Chebyshev, on replacing x by $\frac{1}{2}x$ and subtracting, we have

$$\sum_{n\leqslant x} \ln n - 2 \sum_{n\leqslant x/2} \ln n = x \ln 2 + O(\ln x) = \sum_{n\leqslant x} \varphi\left(\frac{x}{n}\right)$$

$$- 2 \sum_{n\leqslant x} \psi\left(\frac{x}{2n}\right) = \psi(x) - \psi\left(\frac{x}{2}\right) + \psi\left(\frac{x}{3}\right) - \cdots < \psi(x) \tag{3.1.9}$$

in view of the monotonicity of $\psi(x)$. But $\psi(x)$ is directly connected with $\pi(x)$ and $\Pi(x)$,

$$\pi(x) = \sum_{p\leqslant k} 1, \quad \Pi(x) = \sum_{p^k\leqslant x} 1,$$

$$\Pi(x) = \pi(x) + \pi(\sqrt{x}) + \cdots = \pi(x) + O(\sqrt{x}).$$

In fact,

$$\pi(x) = \Pi(x) + O(\sqrt{x}) = \sum_{n\leqslant x} \frac{\Lambda(n)}{\ln n} + O(\sqrt{x})$$

$$= \sum_{n\leqslant x} \frac{\psi(n) - \psi(n-1)}{\ln n} + O(\sqrt{x}) = \frac{\psi(x)}{\ln x}$$

$$+ \sum_{n=1}^{|x|-1} \psi(n) \left[\frac{1}{\ln n} - \frac{1}{\ln(n-1)}\right] + O(\sqrt{x}) = \frac{\psi(x)}{\ln x}$$

$$+ \sum_{n\leqslant x} \frac{\psi(n)}{n \ln^2 n} + O(\sqrt{x}) = \frac{\psi(x)}{\ln x} + \int_2^x \frac{\psi(t)}{t \ln^2 t}\, dt + O(\sqrt{x})$$

$$= \int_2^x \frac{dt}{\ln t} + \frac{\psi(x) - x}{\ln x} + \int_2^x \frac{\psi(t)}{t \ln^2 t}\, dt + O(\sqrt{x}). \tag{3.1.10}$$

We now have from the inequality (3.1.9):

$$\pi(x) > \frac{x}{\ln x} \ln 2 + O\left(\frac{x}{\ln^2 x}\right), \tag{3.1.11}$$

i.e. the elementary Chebyshev lower bound for $\pi(x)$. We give below a proof of the asymptotic law for $\pi(x)$, in other words, $\pi(x) = x/(\ln x) + 0[x/(\ln x)]$; for the present we shall dwell on the question of the number of primes in an arithmetic progression.

2. Dirichlet's Theorem on the Infinity of Primes in an Arithmetic Progression

We shall recall the basic properties of characters modulo $D > 1$. All the characters are multiplicative, $\chi(n)\,\chi(m) = \chi(nm)$, and if $\chi \neq \chi_0$, in other words, the character is non-principal, we have

$$\left.\begin{gathered}\left|\sum_{k=n}^{N}\chi(k)\right| < h,\quad \left|\sum_{k=n}^{N}\chi(k)\,U_k\right| < 2hU_n,\\ U_n \geqslant U_{n+1} \geqslant \cdots \geqslant U_N \geqslant 0,\end{gathered}\right\} \tag{3.2.1}$$

where $h = \varphi(D)$ is Euler's function.

If $\chi(n)$ is a real character, $\chi \neq \chi_0$, then

$$L(\chi) = \sum_{n=1}^{\infty}\frac{\chi(n)}{n} \neq 0. \tag{3.2.2}$$

The convergence of this series for any non-principal character is ensured by inequalities (3.2.1). For,

$$\left|\sum_{n}^{N}\frac{\chi(k)}{k}\right| < \frac{2h}{n},$$

whence also

$$\left|L(\chi) - \sum_{n\leqslant x}\frac{\chi(n)}{n}\right| < \frac{2h}{x},\quad L(\chi) = \sum_{n=1}^{\infty}\frac{\chi(n)}{n}. \tag{3.2.3}$$

To prove (3.2.2) we consider the function $U(x)$ $(0 < x < 1)$:

$$U(x) = \sum_{n=1}^{\infty}\frac{\chi(n)\,x^n}{1-x^n} = \sum_{n=1}^{\infty}\left[\sum_{d|n}\chi(d)\right]x^n = \sum_{n=1}^{\infty}f_n x^n.$$

We remark first of all that

$$f_n = \prod_{k=1}^{s}\left[1 + \chi(p_k) + \cdots + \chi^{\nu_k}(p_k)\right] \qquad (n = p_1^{\nu_1}\cdots p_s^{\nu_s}).$$

Hence it follows, since $\chi(p) = 1,\ -1,\ 0$, that we always have $f_n \geqslant 0$ and $f_n \geqslant 1$, if all the ν_k are even, in other words, if $n = n_0^2$. Hence, with $1 > x > \frac{1}{2}$, $x > x_0$,

$$\begin{aligned}U(x) > \sum_{n=1}^{\infty}x^{n^2} &= \int_1^{\infty}x^{t^2}\,dt + O(1)\\ &= \frac{1}{\sqrt{-\ln x}}\int_0^{\infty}e^{-t^2}\,dt + O(1) = \frac{\sqrt{\pi}}{2\sqrt{-\ln x}} + O(1)\\ &= \frac{\sqrt{\pi}}{2}\,\frac{1}{[-\ln(1-(1-x))]^{1/2}} + O(1) > \frac{1}{2\sqrt{1-x}}.\end{aligned}$$

Further, when

$$S_n = \sum_{k=n}^{\infty} \frac{\chi(k)}{k}, \quad S_1 = L(\chi),$$

$$R_1(x) = U(x) - \frac{L(\chi)}{1-x} = \sum \chi(n) \frac{x^n}{1-x^n} - \sum_{n=1}^{\infty} \frac{\chi(n)}{n} \frac{x^n}{1-x}$$

$$+ \sum_{n=1}^{\infty} (S_n - S_{n+1}) \frac{x^n}{1-x} - \frac{L(\chi)}{1-x}$$

$$= \sum \chi(n) \left[\frac{x^n}{1-x^n} - \frac{x^n}{n(1-x)} \right] - \sum_{n=0}^{\infty} S_{n+1} x^n.$$

By virtue of (3.2.3),

$$\sum_{n=1}^{\infty} S_{n+1} x^n = O\left[\sum_{n=1}^{\infty} \frac{x^n}{n} \right] = O\left(\ln \frac{1}{1-x} \right).$$

Further, we obtain with the aid of Abel's transformation:

$$R_2 = \left| \sum_{n=1}^{\infty} \chi(n) \left[\frac{x^n}{1-x^n} - \frac{x^n}{n(1-x)} \right] \right|$$

$$= \left| \sum_{n=1}^{\infty} T(n) \left[\frac{x^n}{1-x^n} - \frac{x^n}{n(1-x)} - \frac{x^{n+1}}{1-x^{n+1}} + \frac{x^{n+1}}{(n+1)(1-x)} \right] \right|$$

$$< \frac{h}{1-x} \sum_{n=1}^{\infty} \left| \frac{x^n}{1+x+\cdots+x^{n-1}} - \frac{x^{n+1}}{1+x+\cdots+x^n} \right.$$

$$\left. - \frac{x^n}{n(n+1)} - \frac{(1-x)\,x^n}{n+1} \right| < \frac{h}{1-x} \sum_{n=1}^{\infty} \left[\frac{x^n}{1+x+\cdots+x^{n-1}} \right.$$

$$\left. - \frac{x^{n+1}}{1+x+\cdots+x^n} - \frac{x^n}{n(n+1)} \right] + h \sum_{1}^{\infty} \frac{x^n}{n+1}$$

$$= 2h \sum_{1}^{\infty} \frac{x^n}{n+1} = O\left(\ln \frac{1}{1-x} \right),$$

$$T(n) = \sum_{k=1}^{n} \chi(k), \quad |T_k| < h.$$

Thus

$$U(x) = O\left(\ln \frac{1}{1-x} \right), \quad U(x) > \frac{1}{2\sqrt{1-x}},$$

if $L(\chi) = 0$. This contradiction proves (3.2.2). We recall also that

$$\sum_{\chi} \chi(nl') = \begin{cases} h, & n \equiv l \pmod{D} \\ 0, & n \not\equiv l \pmod{D}, \end{cases} \qquad ll' \equiv 1 \pmod{D},$$

where the summation is over all χ, $(l, D) = 1$.

Now let $\chi(n)$ again be a non-principal character. Now,

$$\sum_{n\leqslant x}\frac{\chi(n)\ln n}{n}=\sum_{n\leqslant x}\frac{\chi(n)}{n}\sum_{d|n}\Lambda(d)$$

$$=\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n}\sum_{k\leqslant x/n}\frac{\chi(k)}{k}=L(\chi)\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n}+R(x),$$

where

$$R(x)=\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n}\sum_{k>x/n}\frac{\chi(k)}{k}=\sum\frac{\Lambda(n)}{n}O\left(\frac{n}{x}\right)$$

$$=O\left[\frac{1}{x}\sum_{n\leqslant x}\Lambda(n)\right]=O(1).$$

In addition to this, by (3.2.1), the series $\sum_{n=1}^{\infty}(\chi(n)\ln n)/n$ is convergent.

Hence we obtain the relationship

$$L(\chi)\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n}=O(1). \tag{3.2.4}$$

We must also have for the non-principal character $\chi(n)$:

$$\sum_{n\leqslant x}\frac{\chi(n)}{n}\sum_{d|n}\mu(d)\ln\frac{x}{d}=\ln x+\sum_{n\leqslant 1}\frac{\chi(n)\Lambda(n)}{n}$$

$$=\sum_{n\leqslant x}\frac{\chi(n)\mu(n)}{n}\ln\frac{x}{n}\sum_{k\leqslant x/n}\frac{\chi(k)}{k}=L(\chi)\sum_{n\leqslant x}\frac{\chi(n)\mu(n)}{n}\ln\frac{x}{n}+R(x),$$

where

$$R(x)\sum_{n\leqslant x}\frac{1}{n}\ln\frac{x}{n}O\left(\frac{n}{x}\right)=O\left(\sum_{n\leqslant x}\frac{1}{x}\ln\frac{x}{n}\right)=O(1).$$

Hence

$$\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n}=-\ln x+L(\chi)\,A(x)+O(1). \tag{3.2.5}$$

It follows from (3.2.4) and (3.2.5) that, when $\chi(n)$ is a non-principal character,

$$\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n}=\delta(\chi)\ln x+O(1),$$

$$\delta(\chi)=\begin{cases}-1, & L(\chi)=0,\\ 0, & L(\chi)\neq 0.\end{cases} \tag{3.2.6}$$

If $\chi = \chi_0$ is the principal character, then

$$\sum_{n\leqslant x} \frac{\chi_0(n)\,\Lambda(n)}{n} = \sum_{n\leqslant x} \frac{\Lambda(n)}{n} + O(1) = \ln x + O(1). \quad (3.2.7)$$

We sum over all χ the left- and right-hand sides of (3.2.6) and (3.2.7). We then have

$$\ln x + \Big[\sum_{\chi\neq\chi_0} \delta(\chi)\Big] \ln x + O(1) = \sum \frac{\Lambda(n)}{n} \sum_{\chi} \chi(n) \geqslant 0. \quad (3.2.8)$$

Suppose now that $L(\chi) = 0$ for a character of the third kind, which is not real. Now, $L(\chi) = 0$ also, in other words, in this case $\delta(\chi) = \delta(\overline{\chi}) = -1$. Hence

$$(1-2)\ln x = -\ln x > O(1).$$

This contradiction shows that $L(\chi) \neq 0$ for characters of the third kind. For characters of the second kind $\delta(\chi) = 0$, since $L(\chi) \neq 0$ by virtue of (3.2.2). Hence, for any $\chi \neq \chi_0$,

$$\sum_{n\leqslant x} \frac{\chi(n)\,\Lambda(n)}{n} = O(1). \quad (3.2.9)$$

We shall now prove a theorem which is rather stronger than Dirichlet's theorem on progressions.

THEOREM 3.2.1. *If $D > 1$, l are integers, $(D, l) = 1$, then*

$$\sum_{\substack{p\leqslant x \\ p\equiv l(\mathrm{mod}\ D)}} \frac{\ln p}{p} = \frac{1}{\varphi(D)} \ln x + O(1), \quad (3.2.10)$$

where the p are primes and $\varphi(D)$ is Euler's function.

Let $ll' \equiv 1 \pmod{D}$. We multiply the left-hand sides of (3.2.6) and (3.2.7) by $\chi(l')$, respectively, and sum over all χ. We have

$$\sum_{\chi} \chi(l') \sum_{n\leqslant x} \frac{\chi(n)\,\Lambda(n)}{n} = \ln x + O(1)$$

$$= \varphi(D) \sum_{\substack{n\leqslant x \\ n\equiv l(\mathrm{mod}\ D)}} \frac{\Lambda(n)}{n} + O(1) = \varphi(D) \sum_{\substack{p\leqslant x \\ p\equiv l(\mathrm{mod}\ D)}} \frac{\ln p}{p} + O(1)$$

on the basis of (3.2.3), (3.2.6) and (3.2.7). This proves our theorem.

3. Fundamental Inequalities for Estimating the Number of Primes in the Series of Positive Integers

We now turn to the proof of the fundamental limit theorems on the distribution of primes in the series of positive integers and in progressions. We already know from (3.1.10) that

$$\pi(x) = \frac{\psi(x)}{x} + O\left[\sum_{2}^{x} \frac{1}{\ln^2 n}\right] + O(\sqrt{x}) = \frac{\psi(x)}{\ln x} + O\left(\frac{x}{\ln^2 x}\right),$$

in other words, that $\psi(x) = x + o(x)$ implies

$$\pi(x) = \frac{x}{\ln x} + o\left(\frac{x}{\ln x}\right). \tag{3.3.1}$$

Similarly, having defined, for $D > 1$, $(l, D) = 1$, the functions

$$\pi(l, D; x) = \sum_{\substack{p \leqslant x \\ p \equiv l (\mathrm{mod}\, D)}} 1, \quad \psi(l, x) = \sum_{\substack{n \leqslant x \\ n \equiv l (\mathrm{mod}\, D)}} \Lambda(n),$$

we have

$$\pi(l, D; x) = \sum_{\substack{n \leqslant x \\ n \equiv l (\mathrm{mod}\, D)}} \frac{\Lambda(n)}{\ln n} + O(\sqrt{x}) = \sum_{n \leqslant x} \frac{\psi(l, n) - \psi(l, n-1)}{\ln n}$$

$$+ O(\sqrt{x}) = \frac{\psi(l, x)}{\ln x} + \sum_{n \leqslant x} \frac{\psi(l, n)}{n \ln^2 n} + O(\sqrt{x})$$

$$= \frac{\psi(l, x)}{\ln x} + O\left(\frac{x}{\ln^2 x}\right).$$

Hence it follows at once that, if $\psi(l, x) = x/h + o(x)$, $h = \varphi(D)$, then

$$\pi(l, D, x) = \frac{1}{h} \frac{x}{\ln x} + o\left(\frac{x}{\ln x}\right). \tag{3.3.2}$$

We now define the functions

$$M(x) = \sum_{n \leqslant x} \mu(n), \quad M(l, x) = \sum_{\substack{n \leqslant x \\ n \equiv l (\mathrm{mod}\, D)}} \mu(n), \tag{3.3.3}$$

where $(l, D) = 1$, and show that, instead of estimates of the functions $\psi(x)$ and $\psi(l, x)$, we can use the simpler estimates of the functions $M(x)$ and $M(l, x)$. We shall prove as a preliminary that, when $n > \sqrt{x}$,

$$\Delta = \left[\frac{x}{n}\right] - \left[\frac{x}{n+1}\right] = \begin{cases} 1, & n = \left[\dfrac{x}{k}\right], \\ 0, & n \neq \left[\dfrac{x}{k}\right], \end{cases} \tag{3.3.4}$$

where $k < \sqrt{x}$ is an integer. In fact,

$$\Delta = \left[\frac{x}{n}\right] - \left[\frac{x}{n+1}\right] = \left[\frac{x}{n}\right] - \left[\left[\frac{x}{n}\right] + \left\{\frac{x}{n}\right\} - \frac{x}{n(n+1)}\right]$$

$$= \left[\left\{\frac{x}{n}\right\} - \frac{x}{n(n+1)}\right] = 0,$$

if $\left\{\frac{x}{n}\right\} \geqslant \frac{x}{n(n+1)}$. In other words, $\Delta = 1$ when and only when the reverse inequality $\left\{\frac{x}{n}\right\} < \frac{x}{n(n+1)}$ is satisfied. But in this case, when $k = \left[\frac{x}{n}\right]$, we shall have

$$\frac{n}{k}\left\{\frac{x}{n}\right\} < \frac{n}{k}\frac{x}{n(n+1)} = \frac{x}{(n+1)\frac{x}{n} - (n+1)\left\{\frac{x}{n}\right\}}$$

$$< \frac{x}{(n+1)\frac{x}{n} - \frac{x}{n}} = 1.$$

Thus finally,

$$\frac{x}{k} = n + \frac{n}{k}\left\{\frac{x}{n}\right\} = n + \left\{\frac{x}{k}\right\},$$

or $n = \left[\frac{x}{k}\right]$.

Let $V(n)$ be the number of divisors of the number n. Now, in view of (3.1.2),

$$V(n) = \sum_{d|n} 1, \quad \sum_{d|n} \mu(d)\, V\left(\frac{n}{d}\right) = 1. \tag{3.3.5}$$

Further, we can write

$$T(x) = \sum_{n\leqslant x} V(n) = \sum_{n\leqslant x}\sum_{d|n} 1 = \sum_{n\leqslant x}\sum_{k\leqslant x/n} 1 = \sum_{n\leqslant x}\left[\frac{x}{n}\right]$$

$$= \sum_{n\leqslant\sqrt{x}}\left[\frac{x}{n}\right] + \sum_{\sqrt{x}<n\leqslant x}[n-(n-1)]\left[\frac{x}{n}\right] = \sum_{n\leqslant\sqrt{x}}\left[\frac{x}{n}\right]$$

$$- [\sqrt{x}]\left[\frac{x}{[\sqrt{x}]+1}\right] + \sum_{\sqrt{x}<n\leqslant x}\left(\left[\frac{x}{n}\right] - \left[\frac{x}{n+1}\right]\right)n$$

$$= 2\sum_{n\leqslant\sqrt{x}}\left[\frac{x}{n}\right] + O(\sqrt{x}) - x = 2x\sum_{n\leqslant\sqrt{x}}\frac{1}{n} - x + O(\sqrt{x})$$

$$= x\ln x + (2C-1)x + O(\sqrt{x}) \tag{3.3.6}$$

in view of (3.3.4) and the relationship $\sum_{k\leqslant x} 1/k = \ln x + C + O(1/x)$.

This is the familiar Dirichlet relationship. We shall now prove a lemma.

LEMMA 1. *If* $M(x) = o(x)$, *then*

$$\pi(x) = \frac{x}{\ln x} + o\left(\frac{x}{\ln x}\right). \tag{3.3.7}$$

Whereas, if not only $M(x) = o(x)$, *but also* $M(l, x) = o(x)$ *for all* $l < D$, $(l, D) = 1$, *then also*

$$\pi(l, D; x) = \frac{1}{h}\frac{x}{\ln x} + o\left(\frac{\ln x}{x}\right), \tag{3.3.8}$$

where $(l, D) = 1$ *and* $h = \varphi(D)$ *is Euler's function.*

We consider the relationship, following directly from (3.1.1), (3.1.3) and (3.3.5):

$$\sum_{n\leqslant x}\sum_{d/n}\mu(d)\left[\ln\frac{n}{d} - V\left(\frac{n}{d}\right) + 2C\right] = \psi(x) - [x] + 2C.$$

We obtain from this relationship, on changing the order of summation,

$$\begin{aligned}\psi(x) - [x] + 2C &= \sum_{n\leqslant x}\mu(n)\sum_{k\leqslant x/n}[\ln k - V(k) + 2C] \\ &= \sum_{n\leqslant N}\mu(n)\,W\left(\frac{x}{n}\right) - M(N)\,W\left(\frac{x}{N+1}\right) \\ &\quad + \sum_{n<N\leqslant x} M(n)\left[W\left(\frac{x}{n}\right) - W\left(\frac{x}{n+1}\right)\right],\end{aligned}$$

$$W\left(\frac{x}{n}\right) = \frac{x}{n}\ln\frac{x}{n} - \frac{x}{n} + O\left(\ln\frac{x}{n}\right) - \sum_{k\leqslant x/n} V(k) + 2C\left[\frac{x}{n}\right] = C\left(\sqrt{\frac{x}{n}}\right).$$

for any N, $N = \left[\frac{x}{t}\right]$, $t < \sqrt{x}$. Using (3.3.6) and (3.3.4), we finally obtain the inequality

$$\begin{aligned}|\psi(x) - x| &< C\left[\sum_{n\leqslant N}\sqrt{\frac{x}{n}} + M\left(\frac{x}{t}\right)\sqrt{t}\right. \\ &\quad \left. + \sum_{k\leqslant t}\left|\ln k - V(k) + 2C\right| \max_{\leqslant k\leqslant t}\left|M\left(\frac{x}{k}\right)\right|\right] \\ &< C_1\left[\frac{x}{\sqrt{t}} + 2\max_{1\leqslant k\leqslant t}\left|M\left(\frac{x}{k}\right)\right| t\ln t\right],\end{aligned}$$

where C and C_1 are independent of x and t. This inequality shows that $\psi(x) = x + o(x)$, since, given a fixed t, no matter how large,

and the condition $M(x) = o(x)$, an x can be chosen so large that the inequality

$$\max_{1 \leqslant k \leqslant t} \left| M\left(\frac{x}{k}\right) \right| < \frac{x}{t^2}$$

is satisfied. But, in view of (3.3.1), $\psi(x) = x + o(x)$ implies (3.3.7).

We now observe that, if $\chi_0(n)$ is a principal character modulo D, then

$$\sum_{n \leqslant x} \chi_0(n)\, \Lambda(n) = \sum_{n \leqslant x} \Lambda(n) + O(\ln x) = \psi(x) + O(\ln x), \quad (3.3.9)$$

since the number D is divisible by a bounded number of primes. Further, if $\chi(n)$ is a non-principal character modulo D, then on changing the order of summation, with $N = \left[\frac{x}{t}\right]$,

$$\begin{aligned}
\sum_{n \leqslant x} \chi(n)\, \Lambda(n) &= \sum_{n \leqslant x} \chi(n) \sum_{d \mid n} \mu(d) \ln \frac{n}{d} \\
&= \sum_{n \leqslant x} \chi(n)\, \mu(n) \sum_{k \leqslant x/n} \chi(k) \ln k = \sum_{n \leqslant N} \chi(n)\, \mu(n)\, W\left(\frac{x}{n}\right) \\
&\quad + \sum_{N < n \leqslant x} [M_0(\chi, n) - M_0(\chi, n-1)]\, W\left(\frac{x}{n}\right) \\
&= \sum_{n \leqslant N} \chi(n)\, \mu(n)\, W\left(\frac{x}{n}\right) - M_0(\chi, N)\, W\left(\frac{x}{N+1}\right) \\
&\quad + \sum_{N < n \leqslant x} M_0(\chi, n) \left[W\left(\frac{x}{n}\right) - W\left(\frac{x}{n+1}\right) \right], \quad (3.3.10)
\end{aligned}$$

$$W\left(\frac{x}{n}\right) = \sum_{k \leqslant x/n} \chi(k) \ln k,$$

where

$$M_0(\chi, x) = \sum_{k \leqslant x} \chi(k)\, \mu(k) = \sum_{l=1}^{D-1} \chi(l)\, M(l, x) = o(x)$$

under the assumptions of our lemma. Since

$$\begin{aligned}
W(x) &= \sum_{k \leqslant x} \chi(k) \ln k \\
&= \sum_{k \leqslant x} \chi(k) \ln x \sum_{n \leqslant x-1} \ln\left(1 + \frac{1}{n}\right) \sum_{k \leqslant n} \chi(k) = O(\ln x),
\end{aligned}$$

we find from (3.3.10) with the aid of (3.3.4) that

$$\sum_{n\leqslant x}\chi(n)\,\Lambda(n) < C\left[\sum_{n\leqslant N}\ln\frac{x}{n}+\left|M_0\left(\chi,\frac{x}{t}\right)\right|\ln t\right.$$
$$\left.+\sum_{n\leqslant t}\left|M_0\left(\chi,\frac{x}{n}\right)\right|\ln n\right]$$
$$< C_1\left[\frac{x}{t}\ln t+\max_{n\leqslant t}\left|M\left(\chi,\frac{x}{n}\right)\right|t\ln t\right], \qquad (3.3.11)$$

where C and C_1 are independent of x and t, if $x \geqslant t^2$. This inequality shows that, when $\chi(n)$ is a non-principal character,

$$\sum\chi(n)\,\Lambda(n) = o(x), \qquad (3.3.12)$$

since we can take t as large as desired on the right-hand side of (3.3.11), and x so large that

$$\left|M_0\left(\chi,\frac{x}{n}\right)\right| < \frac{x}{t^2}\,(1\leqslant n\leqslant t),$$

which is possible by virtue of the conditions of our lemma. But now, when $\psi(x) = o(x) + x$,

$$\psi(l,\,x) = \frac{1}{h}\sum_{\chi}\chi(l')\sum_{n\leqslant x}\chi(n)\,\Lambda(n) = \frac{x}{h}+o(x),$$

where $ll' \equiv 1 \pmod{D}$ and $h = \varphi(D)$ is Euler's function. In view of (3.3.2), this last expression leads to the assertion (3.3.8) of our lemma. Lemma 1 reduces asymptotic formulae for the functions $\pi(x)$ and $\pi(l, D, x)$ to asymptotic formulae for functions $M(x)$ and $M(l, x)$. For these latter formulae, we shall require some new relationships, and some asymptotic formulae of a general nature.

We observe first of all that

$$\sum_{d/n}\mu(d)\ln^2 d = \frac{d^2}{ds^2}\left[\sum_{d/n}(d)\,d^s\right]\Big|_{s=0}$$
$$= \frac{d^2s}{ds^2}\left[\prod_{p/n}(1-p^s)\right]_{s=0} = \begin{cases} -\ln^2 p, & n = p^k, \\ 2\ln p\ln q, & n = p^kq^r, \\ 0, & p_1p_2p_3/n, \end{cases} \qquad (3.3.13)$$

where p, q, p_1, p_2, p_3 are different primes. Hence

$$\sum_{n\leqslant x}\sum_{d|n}\mu(d)\ln^2\frac{x}{d} = \ln^2 x + 2\ln x\sum_{n\leqslant x}\Lambda(n) + \sum_{nm\leqslant x}\Lambda(n)\Lambda(m) - \sum_{n\leqslant x}\Lambda(n)\ln n = 2\ln x\psi(x) - \ln x\psi(x) + \sum_{nm\leqslant x}\Lambda(n)\Lambda(m) + O(x)$$

$$= \psi(x)\ln x + \sum_{n\leqslant x}\Lambda(n)\,\psi\left(\frac{x}{n}\right) + O(x) = \psi(x)\ln x + \sum_{nm\leqslant x}\Lambda(n)\Lambda(m) + O(x), \tag{3.3.14}$$

since

$$\sum_{n\leqslant x}\Lambda(n)\ln n = \sum_{n\leqslant x}[\psi(n) - \psi(n-1)]\ln n = \psi(x)\ln x - \sum_{n\leqslant x-1}\psi(n)\ln\left(1+\frac{1}{n}\right) + O(1) = \psi(x)\ln x + O(x)$$

by virtue of (3.1.6). Further,

$$\sum_{n\leqslant x}\sum_{d|n}\mu(d)\ln^2\frac{x}{d} = \sum_{n\leqslant x}\mu(n)\ln^2\frac{x}{n}\left[\frac{x}{n}\right] = x\sum_{n\leqslant x}\frac{\mu(n)}{n}\ln^2\frac{x}{n} - \sum_{n\leqslant x}\mu(n)\left\{\frac{x}{n}\right\}\ln^2\frac{x}{n}.$$

We shall now make use of the relationship

$$\sum_{n\leqslant x}\frac{1}{n}\sum_{d|n}\mu(d)\ln\frac{x}{d} = \ln x + \sum_{n\leqslant x}\frac{\Lambda(n)}{n} = 2\ln x + O(1)$$

$$= \sum_{n\leqslant x}\frac{\mu(n)}{n}\ln\frac{x}{n}\sum_{k\leqslant x/n}\frac{1}{k} = \sum_{n\leqslant x}\frac{\mu(n)}{n}\ln\frac{x}{n}\left[\ln\frac{x}{n} + C + O\left(\frac{n}{x}\right)\right]$$

$$= \sum_{n\leqslant x}\frac{\mu(n)}{n}\ln^2\frac{x}{n} + C\sum_{n\leqslant x}\frac{\mu(n)}{n}\left[\sum_{k\leqslant x/n}\frac{1}{k} - C + O\left(\frac{n}{x}\right)\right] + O\left[\frac{1}{x}\sum_{n\leqslant x}\ln\frac{x}{n}\right] = \sum_{n\leqslant x}\frac{\mu(n)}{n}\ln^2\frac{x}{n} + O(1),$$

which holds by virtue of (3.1.2), (3.1.4), (3.1.5), (3.1.6) and the inequality

$$\left|\sum_{n\leqslant x}\mu(n)\left\{\frac{x}{n}\right\}\ln^2\frac{x}{n}\right| < \sum_{n\leqslant x}\ln^2\frac{x}{n} = O(x).$$

We now have the formula

$$\sum_{n\leq x}\sum_{d|n}\mu(d)\ln^2\frac{x}{d}=2x\ln x+O(x), \tag{3.3.15}$$

whence it follows that

$$\psi_1(x)=\sum_{nm\leq x}\Lambda(n)\Lambda(m)=2x\ln x-\psi(x)\ln x+O(x). \tag{3.3.16}$$

We now introduce the function $\psi_2(x)$:

$$\begin{aligned}\psi_2(x)&=\sum_{nm\leq x}\frac{\Lambda(n)\Lambda(m)}{\ln nm}=\sum_{n\leq x}\frac{\psi_1(n)-\psi_1(n-1)}{\ln n}\\&=\frac{\psi_1(x)}{\ln x}+\sum_{k\leq x-1}\psi(k)\left[\frac{1}{\ln k}-\frac{1}{\ln(k+1)}\right]+O(\ln k)\\&=2x-\psi(x)+O\left(\frac{x}{\ln x}\right)+\sum_{n\leq x-1}\frac{2x\ln n-\psi(n)\ln n+O(n)}{n\ln^2(n+1)}\\&=2x-\psi(x)+O\left(\frac{x}{\ln x}\right).\end{aligned} \tag{3.3.17}$$

Now,

$$\begin{aligned}\psi_2(x)+\psi(x)&=\sum_{n\leq x}\Lambda(n)+\sum_{nm\leq x}\frac{\Lambda(n)\Lambda(m)}{\ln nm}\\&=2x+O\left(\frac{x}{\ln x}\right).\end{aligned} \tag{3.3.18}$$

We next consider the relationship

$$\begin{aligned}R_l(x)&=\sum_{\substack{n\equiv l(\mathrm{mod}\,D)\\ n\leq x}}\sum_{d|n}\mu(d)\ln^2\frac{x}{d}=2\ln x\sum_{\substack{n\leq x\\ n\equiv l(\mathrm{mod}\,D)}}\Lambda(n)\\&\quad-\sum_{\substack{n\leq x\\ n\equiv l(\mathrm{mod}\,D)}}\Lambda(n)\ln n+\sum_{\substack{nm\leq x\\ nm\equiv l(\mathrm{mod}\,D)}}\Lambda(n)\Lambda(m)+O(\ln x)\\&=\psi(l,x)\ln x+\sum_{\substack{mn\leq x\\ mn\equiv l(\mathrm{mod}\,D)}}\Lambda(n)\Lambda(m)+O(x),\end{aligned}$$

which holds by virtue of the fact that

$$\begin{aligned}\sum_{\substack{n\leq x\\ n\equiv l(\mathrm{mod}\,D)}}\Lambda(n)\ln n&=\psi(l,x)\ln x-\sum_{n\leq x}\frac{1}{n}\,\frac{\psi(l,n)}{n}+O(1)\\&=\psi(l,x)\ln x+O(x).\end{aligned}$$

We observe first of all that, when $q > 1$,

$$\sum_{\substack{q|n \\ n\leqslant x}} \sum_{d|n} \mu(d) \ln^2 \frac{x}{d} = 2 \ln x \sum_{\substack{n\leqslant x \\ d|n}} \Lambda(n) + O(\ln x)$$
$$- \sum_{\substack{n\leqslant x \\ q|n}} \Lambda(n) \ln n + \sum_{\substack{mn\leqslant x \\ q|mn}} \Lambda(n)\,\Lambda(m) = O(\ln^2 x)$$
$$+ O\left[\sum_{n\leqslant x} \Lambda(n) \ln \frac{x}{n}\right] + O\left[\sum_{p\leqslant x} \sum_{k=2}^{\infty} \ln p \left[\frac{x}{p^k}\right]\right] = O(x),$$
$$\sum_{n<x} \Lambda(n) \sum_{\substack{mn\leqslant x \\ q|mn}} \Lambda(m) = \sum_{\substack{mn\leqslant x \\ q|mn}} \Lambda(n)\,\Lambda(m).$$

Hence, if χ_0 is a principal character modulo D, we have, with $D = p_1^{\alpha_1} \dots p_\nu^{\alpha_\nu}$,

$$\sum_{n\leqslant x} \chi_0(n) \sum_{d|n} \mu(d) \ln^2 \frac{x}{d} = \sum_{m|D} \mu(m) \sum_{\substack{n\leqslant x \\ m|n}} \sum_{d|n} \mu(d) \ln^2 \frac{x}{d}$$
$$= \sum_{n\leqslant x} \sum_{d|n} \ln^2 \frac{x}{d} + O(x) = 2x \ln x + O(x)$$

by virtue of (3.3.15). Further, if $\chi(n)$ is a non-principal character modulo D, then

$$\sum_{x\leqslant n} \chi(n) \sum_{d|n} \mu(d) \ln^2 \frac{x}{d} = \sum_{d\leqslant x} \mu(d) \ln^2 \frac{x}{d} \sum_{k\leqslant x/d} \chi(k)$$
$$= O\left(\sum_{n\leqslant x} \ln^2 \frac{x}{n}\right) = O(x).$$

Hence we obtain finally, in the case when $l'l \equiv 1 \pmod D$,

$$R_l(x) = \frac{1}{\varphi(D)} \sum_{\chi} \chi(l') \sum_{n\leqslant x} \chi(n) \sum_{d|n} \mu(d) \ln^2 \frac{x}{d} = \frac{2x \ln x}{\varphi(D)} + O(x).$$

In other words, we have

$$\sum_{\substack{nm<x \\ nm\equiv l(\mathrm{mod}\, D)}} \Lambda(n)\,\Lambda(m) = \psi_1(l, x) = \frac{2}{h} x \ln x - \psi(l, x) \ln x + O(x),$$
$$h = \varphi(D). \tag{3.3.19}$$

On using these relationships, we also find that

$$\psi_2(l, x) = \sum_{\substack{nm\leqslant x \\ nm\equiv l(\mathrm{mod}\, D)}} \frac{\Lambda(n)\,\Lambda(m)}{\ln nm} = \sum_{n\leqslant x} \frac{\psi_1(l, n) - \psi_1(l, n-1)}{\ln n}$$
$$= \frac{\psi_1(l, x)}{\ln x} + \sum_{n\leqslant x} \frac{\psi_1(l, n)}{n \ln^2 n} + O(1)$$
$$= -\psi(l, x) + \frac{2}{\varphi(D)} x + O\left(\frac{x}{\ln x}\right), \tag{3.3.20}$$

whence it follows that

$$\psi(l, x) + \psi_2(l, x) = \frac{2}{\varphi(D)} x + O\left(\frac{x}{\ln x}\right). \tag{3.3.21}$$

Further, we have

$$\begin{aligned}
\sum_{n\leqslant x} \mu(n) \ln \frac{x}{n} &= O\left(\sum_{n\leqslant x} \ln \frac{x}{n}\right) = O(x) = M(x) \ln x \\
&\quad - \sum_{n\leqslant x} \mu(n) \sum_{d|n} \Lambda(d) = M(x) \ln x - \sum_{n\leqslant x} \Lambda(n) \sum_{k\leqslant x/n} \mu(nk) \\
&= M(x) \ln x - \sum_{p\leqslant x} \ln p \sum_{k\leqslant x/p} \mu(p \cdot k) \\
&= M(x) \ln x + \sum_{p\leqslant x} \ln p M\left(\frac{x}{p}\right) - \sum_{p\leqslant x} \ln p \sum_{k\leqslant x/p^s} \mu(k) \\
&= M(x) \ln x + \sum_{n\leqslant x} \Lambda(n) M\left(\frac{x}{n}\right) + O(x),
\end{aligned}$$

or

$$\begin{aligned}
M(x) \ln x &= - \sum_{n\leqslant x} \Lambda(n) M\left(\frac{x}{n}\right) + O(x) \\
&= - \sum_{n\leqslant x} \mu(n) \psi\left(\frac{x}{n}\right) + O(x).
\end{aligned} \tag{3.3.22}$$

On replacing $\psi\left(\frac{x}{n}\right)$ by its expression from (3.3.18) in the right-hand side of (3.3.22), we obtain

$$\begin{aligned}
M(x) \ln x &= \sum_{n\leqslant x} \mu(n) \left[\sum_{km\leqslant x/n} \frac{\Lambda(k)\Lambda(m)}{\ln km} - \frac{2x}{n} \right. \\
&\quad \left. + O\left(\frac{x}{n\left[\ln \frac{x}{n} + 1\right]}\right)\right] = -2x \sum_{n\leqslant x} \frac{\mu(n)}{n} \\
&\quad + \sum_{kmn\leqslant x} \frac{\Lambda(k)\Lambda(n)\mu(m)}{\ln km} + O\left[\sum_{n\leqslant x} \frac{x}{n\left(1 + \ln \frac{x}{n}\right)}\right] \\
&= 2xO(1) + \sum_{kn\leqslant x} \frac{\Lambda(k)\Lambda(n)}{\ln kn} M\left(\frac{x}{kn}\right) + O(x \ln \ln x) \\
&= \sum_{kn\leqslant x} \frac{\Lambda(k)\Lambda(n)}{\ln kn} M\left(\frac{x}{kn}\right) + O(x \ln \ln x),
\end{aligned}$$

since

$$\sum_{n\leqslant x}\frac{x}{n\left(1+\ln\frac{x}{n}\right)}=\int_1^x\frac{x\,dt}{t\left(1+\ln\frac{x}{t}\right)}+O(x)$$

$$=x\int_1^x\frac{dt}{t(1+\ln t)}+O(x)=x\ln\ln x+O(x). \qquad (3.3.23)$$

Hence we have the inequality

$$|M(x)|\ln x\leqslant\sum_{kn\leqslant x}\frac{\Lambda(k)\,\Lambda(n)}{\ln kn}\left|M\left(\frac{x}{kn}\right)\right|+O(x\ln\ln x).$$

On adding this inequality term by term to the inequality

$$|M(x)|\ln x\leqslant\sum_{n\leqslant x}\Lambda(n)\left|M\left(\frac{x}{n}\right)\right|+O(x\ln\ln x),$$

which follows at once from (3.3.22), we find that, in view of (3.3.18):

$$\begin{aligned}2|M(x)|\ln x&\leqslant\sum_{n\leqslant x}\Lambda(n)\left|M\left(\frac{x}{n}\right)\right|\\&\quad+\sum_{mn\leqslant x}\frac{\Lambda(n)\,\Lambda(m)}{\ln mn}\left|M\left(\frac{x}{mn}\right)\right|+O(x\ln\ln x)\\&=\sum_{n\leqslant x}[\psi(n)+\psi_2(n)-\psi(n-1)-\psi_2(n-1)]\left|M\left(\frac{x}{n}\right)\right|\\&\quad+O(x\ln\ln x)<2\sum_{n\leqslant x}n\left[\left|M\left(\frac{x}{n}\right)\right|-\left|M\left(\frac{x}{n+1}\right)\right|\right]\\&\quad+O\left[\sum_{n\leqslant x}\frac{n}{1+\ln n}\left|M\left(\frac{x}{n}\right)-M\left(\frac{x}{n+1}\right)\right|\right]+O(x\ln\ln x)\\&=2\sum_{n\leqslant x}\left|M\left(\frac{x}{n}\right)\right|+O\left[\sum_{n\leqslant x}\frac{n}{1+\ln n}\,\frac{x}{n(n+1)}\right]+O(x\ln\ln x)\\&=2\sum_{n\leqslant x}\left|M\left(\frac{x}{n}\right)\right|+O(x\ln\ln x).\end{aligned}$$

We have thus obtained the inequality

$$|M(x)|\ln x\leqslant\sum_{n\leqslant x}\left|M\left(\frac{x}{n}\right)\right|+O(x\ln\ln x). \qquad (3.3.24)$$

We shall next obtain similar inequalities for $|M(l,x)|$.

4. Fundamental Inequalities for Estimating the Number of Primes in a Progression

When we proved Theorem 3.2.1, we also proved in passing that

$$L(\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n} = \sum_{n=1}^{N} \frac{\chi(n)}{n} + O\left(\frac{1}{N}\right) \neq 0 \tag{3.4.1}$$

for any non-principal character modulo D. On using this formula, we obtain

$$\begin{aligned} 1 &= \sum_{n\leqslant x} \frac{\chi(n)}{n} \sum_{d\mid n} \mu(d) = \sum_{n\leqslant x} \frac{\chi(n)\,\mu(n)}{n} \sum_{k\leqslant x/n} \frac{\chi(k)}{k} \\ &= \sum_{n\leqslant x} \frac{\chi(n)\,\mu(n)}{n} \left[L(\chi) + O\left(\frac{n}{x}\right)\right] + O(1) \\ &= L(\chi) \sum_{n\leqslant x} \frac{\chi(n)\,\mu(n)}{n} + O(1), \end{aligned}$$

whence it follows that

$$\sum_{n\leqslant x} \frac{\chi(n)\,\mu(n)}{n} = O(1) \tag{3.4.2}$$

for any non-principal character modulo D. If $\chi_0(n)$ is a principal character modulo D, then

$$\begin{aligned} 1 &= \sum_{n\leqslant x} \chi_0(n) \sum_{d\mid n} \mu(d) = \sum_{n\leqslant x} \chi_0(n)\,\mu(n) \sum_{k\leqslant x/n} \chi_0(k) \\ &= \sum_{n\leqslant x} \chi_0(n)\,\mu(n) \sum_{d\mid D} \mu(d) \left[\frac{x}{nd}\right] \\ &= \sum_{d\mid D} \mu(d) \sum_{n\leqslant x} \chi_0(n)\,\mu(n) \left[\frac{x}{nd} - \left\{\frac{x}{nd}\right\}\right] \\ &= x \sum_{d\mid D} \frac{\mu(d)}{d} \sum_{n\leqslant x} \frac{\chi_0(n)\,\mu(n)}{n} + O(x) \\ &= \frac{\varphi(D)}{D}\, x \sum_{n\leqslant x} \frac{\chi_0(n)\,\mu(n)}{n} + O(x), \end{aligned}$$

whence it follows that

$$\sum_{n\leqslant x} \frac{\chi_0(n)\,\mu(n)}{n} = O(1), \tag{3.4.3}$$

since $\varphi(D) \neq 0$.

Further, we also find from (3.4.2) and (3.4.3) that

$$\sum_{\substack{n\leqslant x\\ n\equiv l(\mathrm{mod}\,D)}} \frac{\mu(n)}{n} = \frac{1}{\varphi(D)} \sum_{\chi} \chi(l') \sum_{n\leqslant x} \frac{\chi(n)\,\mu(n)}{n} = O(1), \qquad (3.4.4)$$

where $(l, D) = 1$, $ll' \equiv 1 \pmod{D}$ and the summation is over all characters modulo D.

We now consider the relationship, when $(l, D) = 1$:

$$\begin{aligned}
\sum_{\substack{n\equiv l(\mathrm{mod}\,D)\\ n\leqslant x}} \mu(n) \ln\frac{x}{n} &= M(l, x)\ln x - \sum_{\substack{n\leqslant x\\ n\equiv l(\mathrm{mod}\,D)}} \mu(n) \sum_{d\mid n} \Lambda(d)\\
&= M(l, x)\ln x - \sum_{n\leqslant x} \Lambda(n) \sum_{\substack{kn\leqslant x\\ kn\equiv l(\mathrm{mod}\,D)}} \mu(kn)\\
&= M(l, x)\ln x - \sum_{p\leqslant x} \ln p \sum_{\substack{pn\leqslant x\\ pn\equiv l(\mathrm{mod}\,D)}} \mu(pn)\\
&= M(l, x)\ln x + \sum_{\substack{p\leqslant x\\ (p, D)=1}} \ln p\, M\left(p'l, \frac{x}{p}\right) - \sum_{\substack{p\leqslant x\\ (p, D)=1}}\\
&\quad \times \ln p \sum_{\substack{p^2k\equiv l(\mathrm{mod}\,D)\\ p^2k\leqslant x}} \mu(pk) = M(l, x)\ln x\\
&\quad + \sum_{\substack{n\leqslant x\\ (n, D)=1}} \Lambda(n)\, M\left(n'l, \frac{x}{n}\right) + O(x),
\end{aligned}$$

$$pp' \equiv 1 (\mathrm{mod}\, D), \quad nn' \equiv 1 (\mathrm{mod}\, D),$$

whence it follows at once that

$$\begin{aligned}
&M(l, x)\ln x + \sum_{n\leqslant x} \chi_0(n)\,\Lambda(n)\, M\left(n'l, \frac{x}{n}\right)\\
&\quad = M(l, x)\ln x + \sum_{n\leqslant x} \mu(n) \sum_{\substack{kn\leqslant x\\ kn\equiv l(\mathrm{mod}\,D)}} \Lambda(k) + O(x)\\
&\quad = M(l, x)\ln x + \sum_{n\leqslant x} \chi_0(n)\,\mu(n)\,\psi\left(n'l, \frac{x}{n}\right) + O(x), \qquad (3.4.5)
\end{aligned}$$

where again $nn' \equiv 1 \pmod{D}$. The first implication of this relationship is that

$$|M(l, x)|\ln x \leqslant \sum_{n\leqslant x} \Lambda(n) \left|\mu\left(n'l, \frac{x}{n}\right)\right| + O(x), \qquad (3.4.6)$$

where again $nn' \equiv 1 \pmod{D}$. On now substituting the expression of (3.3.21) in place of $\psi(n'l, x/n)$ in the right-hand side of (3.4.5),

we obtain

$$M(l, x)\ln x + \sum_{n\leqslant x}\chi_0(n)\,\mu(n)\left[\frac{2}{\varphi(D)}\,\frac{x}{n} + O\left(\frac{x}{n\left(1+\ln\frac{x}{n}\right)}\right)\right]$$

$$- \sum_{n\leqslant x}\chi_0(n)\,\mu(n) \sum_{\substack{km\leqslant x\\ nkm\equiv l(\mathrm{mod}\,D)}} \frac{\Lambda(k)\Lambda(m)}{\ln km} + O(x) = M(l, x)\ln x$$

$$- \sum_{\substack{kmn\leqslant x\\ kmn\equiv l(\mathrm{mod}\,D)}} \frac{\Lambda(k)\,\Lambda(m)\,\mu(n)}{\ln km} + \frac{2x}{\varphi(D)} \sum_{n\leqslant x}\frac{\chi_0(n)\,\mu(n)}{n}$$

$$+ O(x\ln\ln x) = M(l, x)\ln x - \sum_{\substack{km\leqslant x\\ (km,\,D)=1}} \frac{\Lambda(k)\,\Lambda(m)}{\ln km}$$

$$\times M\left((km)'\,l, \frac{x}{n}\right) + O(x\ln\ln x),$$

whence it follows that

$$|M(l, x)|\ln x \leqslant \sum_{km\leqslant x}\chi_0(km)\frac{\Lambda(k)\,\Lambda(m)}{\ln km}\left|M(km)'\,l, \frac{x}{n}\right)\Big| + O(x\ln\ln x), \tag{3.4.7}$$

where $(km)'\,km \equiv 1 \pmod D$. On adding (3.4.7) and (3.4.5), we now find that

$$2|M(l, x)|\ln x \leqslant \sum_{n\leqslant x}\chi_0(n)\,\Lambda(n)\left|M\left(n'l, \frac{x}{n}\right)\right|$$

$$+ \sum_{mn\leqslant x}\chi_0(mn)\frac{\Lambda(n)\,\Lambda(m)}{\ln mn}\left|M\left((nm)'l, \frac{x}{n}\right)\right| + O\,(x\ln\ln x), \tag{3.4.8}$$

where again $nn' \equiv 1 \pmod D$, $(nm)'\,mn \equiv 1 \pmod D$. It further follows from this that

$$2|M(l, x)|\ln x < \sum_{q=1}^{D}\chi_0(q) \sum_{\substack{n\equiv q(\mathrm{mod}\,D)\\ n\leqslant x}} \Lambda(n)\left|M\left(q'l, \frac{x}{n}\right)\right|$$

$$+ \sum_{q=1}^{D}\chi_0(q) \sum_{\substack{mn\equiv q(\mathrm{mod}\,D)\\ mn\leqslant x}} \frac{\Lambda(n)\,\Lambda(m)}{\ln nm}\left|M\left(q'l, \frac{x}{mn}\right)\right| + O(x\ln\ln x)$$

$$= \sum_{q=1}^{D}\chi_0(q) \sum_{\substack{n\equiv q(\mathrm{mod}\,D)\\ n\leqslant x}} [\varphi(q,n) + \varphi_2(q,n) - \varphi(q,n-D)$$

$$- \varphi_2(q,n-D)]\left| M\left(q'l, \frac{x}{n}\right)\right| + O(x \ln\ln x) \leqslant \frac{2D}{\varphi(D)}$$

$$\times \sum_{q=1}^{D} \chi_0(q) \sum_{\substack{n\equiv q(\mathrm{mod}\, D)\\ n<x}} \left| M\left(q'l, \frac{x}{n}\right)\right| + \frac{2D}{\varphi(D)} \sum_{1}^{D} \chi_0(q)$$

$$\times \sum_{\substack{n\equiv q(\mathrm{mod}\, D)\\ n<x}} O\left(\frac{n}{1+\ln n}\right)\left|\left| M\left(q'l, \frac{x}{n}\right)\right| + \left| M\left(q'l, \frac{x}{n+D}\right)\right|\right|$$

$$+ O(x \ln\ln x)$$

by virtue of (3.3.21), where $qq' \equiv 1 \pmod{D}$. But since

$$\sum_{q=1}^{D} \chi_0(q) \sum_{\substack{n\equiv q(\mathrm{mod}\, D)\\ n<x}} O\left(\frac{n}{1+\ln n}\right)\left|\left| M\left(q'l, \frac{x}{n}\right)\right| - \left| M\left(q'l, \frac{x}{n+D}\right)\right|\right|$$

$$< \sum_{1}^{D} \chi_0(q) \sum_{\substack{n\equiv q(\mathrm{mod}\, D)\\ n<x}} O\left(\frac{n}{1+\ln n}\right) \frac{Dx}{n(n+D)} = O\left(x \sum_{n<x} \frac{1}{n \ln n}\right)$$

$$= O(x \ln\ln x),$$

we have:

$$|M(l,x)| \ln x \leqslant \frac{D}{\varphi(D)} \sum_{1}^{D} \chi_0(q) \sum_{\substack{n<x\\ n\equiv q(\mathrm{mod}\, D)}} \left| M\left(q'l, \frac{x}{n}\right)\right|$$

$$+ O(x \ln\ln x).$$

But it is easily seen that, when $(l, D) = 1$,

$$\left| D \sum_{\substack{n\equiv q(\mathrm{mod}\, D)\\ n<x}} \left| M\left(l, \frac{x}{n}\right)\right| - \sum_{n<x} \left| M\left(l, \frac{x}{n}\right)\right|\right|$$

$$\leqslant \left| \sum_{\substack{n\equiv q(\mathrm{mod}\, D)\\ n<x}} \sum_{k=0}^{D-1} \left[\left| M\left(l, \frac{x}{n}\right)\right| - \left| M\left(l, \frac{x}{n+k}\right)\right| \right] + O(x) \right|$$

$$= O\left[\sum_{n<x} \frac{x}{n(n+1)}\right] + O(x) = O(x).$$

Hence we finally have, for any l, $(l, D) = 1$,

$$\left|M(l,x)\right| \ln x \leqslant \frac{1}{\varphi(D)} \sum_{q=1}^{D} \chi_0(q) \sum_{n<x} \left| M\left(q, \frac{x}{n}\right)\right| + O(x) \ln\ln x). \tag{3.4.9}$$

5. Proof of the Limit Theorems for the Distribution of Primes in the Natural Series and in Progressions

Inequalities (3.3.24) and (3.4.9) enable us to show that $M(x) = o(x)$ and $M(l, x) = o(x)$ for all l and D, provided we also make use of arguments that take account of the oscillations of the functions $M(x)$ and $M(l, x)$. First of all, it follows from (3.1.4) that

$$\left| \sum_{n \leqslant x} \frac{M(n) - M(n-1)}{n} \right| = \left| \sum_{n \leqslant x}' \frac{M(n)}{n(n+1)} + \frac{M(x)}{x} \right| < 1,$$

or that

$$\left| \sum_{y \leqslant n \leqslant x} \frac{M(n)}{n(n+1)} \right| < 4. \tag{3.5.1}$$

Similarly, it follows from (3.4.2) and (3.4.3) that, when $(l, D) = 1$,

$$\sum_{\substack{n \leqslant x \\ n \equiv l \pmod D}} \frac{\mu(n)}{n} = \frac{1}{\varphi(D)} \sum_{\chi}' \chi(l') \sum_{n \leqslant x} \frac{\chi(n)\,\mu(n)}{n} = O(1),$$

where $ll' \equiv 1 \pmod D$ and the summation on the right-hand side is over all the characters modulo D. This relationship leads in turn to the asymptotic formula

$$\left| \sum_{y \leqslant n \leqslant x} \frac{M(l, n)}{n(n+1)} \right| = \left| \frac{M(l, y)}{y} - \frac{M(l, x)}{x+1} - \sum_{\substack{n \leqslant x \\ n \equiv l \pmod D}} \frac{\mu(n)}{n} + \sum_{\substack{n \leqslant y \\ n \equiv l \pmod D}} \frac{\mu(n)}{n} \right| < C, \tag{3.5.2}$$

where $C = C(D)$ is a constant, independent of x and l.

Inequalities (3.5.1) and (3.5.2) lead us to the following lemma.

LEMMA 2. *If $x > t > 1$, there exists a point q in the interval $[x, tx]$ such that*

$$|M(q)| < \frac{5q}{\ln t}. \tag{3.5.3}$$

Similarly, if $x > t > 1$, there exists a point q in the interval $[x, tx]$ such that

$$|M(l, q)| < \frac{Cq}{\ln t}, \quad C = C(D), \tag{3.5.4}$$

where C is independent of l, t and x, $(l, D) = 1$.

If $M(x)$ changes sign in $[x, tx]$, there is a point q at which either $M(q) = 0$ or $M(q)\,M(q+1) < 0$. But the second case is impossible, since $M(q)$ differs from $M(q+1)$ by not more than unity and both are integers. Hence, if $M(q)$ changes sign in $[x, tx]$, there exists a point q at which $M(q) = 0$. If $M(x)$ does not change sign in $[x, tx]$, then we find from (3.5.1) that

$$\min_{x\leqslant y\leqslant tx}\left|\frac{M(y)}{y}\right|\sum_{x\leqslant k\leqslant tx}\frac{1}{k}\leqslant\min_{x\leqslant y\leqslant tx}\frac{|M(y)|}{y}\left[\ln t-\frac{1}{x}\right]<4,$$

whence (3.5.3) follows. Further, if $M(l, x)$ changes sign in $[x, tx]$, then just as in the previous case, there is a q in $[x, tx]$ such that $M(l, q) = 0$. If $M(l, q)$ does not change sign in $[x, tx]$, then, again from (3.5.2),

$$\min_{x\leqslant y\leqslant tx}\frac{|M(l,y)|}{y}\sum_{x\leqslant k\leqslant tx}\frac{1}{k}<\min_{x\leqslant y\leqslant tx}\frac{|M(l,y)|}{y}\left[\ln t-\frac{1}{x}\right]<C_0,$$

where C_0 depends only on D. This inequality is equivalent to (3.5.4).

LEMMA 3. *The asymptotic formula* $M(x) = o(x)$ *holds.*

On putting $U(x) = 1/x\,|M(x)|$, we obtain from (3.4.24), on dividing both sides by $x \ln x$:

$$U(x)\leqslant\frac{1}{\ln x}\sum_{n\leqslant x}\frac{1}{n}\,U\left(\frac{x}{n}\right)+O\left(\frac{\ln\ln x}{\ln x}\right). \tag{3.5.5}$$

Let

$$\theta=\overline{\lim_{x\to\infty}}\,U(x)>0.$$

Now, given any ε, $\theta > \varepsilon > 0$ and $x > x_0(\varepsilon)$, $U(x) < \theta + \varepsilon$. Having fixed ε, we select a t subject only to the conditions

$$t>e^6,\qquad 20<\theta\ln t. \tag{3.5.6}$$

By Lemma 2, we can say that there exists a q_k, $t^{2k}\leqslant q_k\leqslant t^{2k+1}$, for which $|M(q_k)| < 5q_k/(\ln t)$. But now,

$$|M(q_k+v)|<|M(q_k)|+v<\frac{5q_k}{\ln t}+v<\frac{10q_k}{\ln t},$$

if $v \ln t < 5q_k$. Hence, if $0 \leqslant v \leqslant [5q_k/\ln t]$, we have

$$U(q_k+v)<\frac{10}{\ln t}.$$

But the numbers n, satisfying the inequalities

$$q_k \leqslant \frac{x}{n} \leqslant q_k + \frac{5q_k}{\ln t} = \left(1 + \frac{5}{\ln t}\right) q_k, \tag{3.5.7}$$

are determined by the inequalities

$$p_k = \left[\frac{1}{1 + \dfrac{5}{\ln t}} \frac{x}{q_k}\right] \leqslant n \leqslant \left[\frac{x}{q_k}\right] = p_k' < p_{k+1}. \tag{3.5.8}$$

Further,

$$\sum_{p_k \leqslant n \leqslant p_k'} \frac{1}{n} = \ln \frac{p_k'}{p_k} + \alpha_k \frac{q_k}{x} = \ln\left(1 + \frac{5}{\ln t}\right) + \beta_k \frac{q_k}{x}, \quad |\beta_k| \leqslant 3. \tag{3.5.9}$$

In other words, when n satisfies (3.5.8), we have $U(x/n) \leqslant 10/(\ln t)$.

On observing that, when $m = \left[\dfrac{1}{2} \dfrac{\ln x}{\ln t} - \dfrac{1}{2}\right]$,

$$\begin{aligned}
\sum_{k=1}^{m} \sum_{p_k \leqslant n \leqslant p_k'} \frac{1}{n} &= m \ln\left(1 + \frac{5}{\ln t}\right) + \sum_{1}^{m} \beta_k \frac{q_k}{x} \\
&= \frac{2 \ln t}{\ln x} \ln\left(1 + \frac{5}{\ln t}\right) + O\left(\sum_{1}^{m} |\beta_k| \frac{t^{2k}}{x}\right) + O(1) \\
&= \lambda \ln x + O(1)\ (\lambda > 0),
\end{aligned} \tag{3.5.10}$$

we at once obtain from (3.5.5), on replacing $U(x/n)$ by $10/(\ln t)$ for all n satisfying (3.5.8), and by $\theta + \varepsilon$ for the remaining n,

$$\begin{aligned}
U(x) &\leqslant \frac{\theta + \varepsilon}{\ln x} \left[\sum_{n \leqslant x} \frac{1}{n} - \sum_{k=1}^{m} \sum_{p_k \leqslant n \leqslant p_k'} \frac{1}{n}\right] \\
&\quad + \frac{10}{\ln t \ln x} \sum_{1}^{m} \sum_{p_k \leqslant n \leqslant p_k'} \frac{1}{n} + O\left(\frac{\ln \ln x}{\ln x}\right) \\
&\leqslant (\theta + \varepsilon)(1 - \lambda) + \frac{10\lambda}{\ln t} + O\left(\frac{\ln \ln x}{\ln x}\right).
\end{aligned}$$

An immediate consequence is an inequality that holds for any ε and $t > D_0$, which is obtained by passing to the upper limit on the left-hand side:

$$\theta \leqslant (\theta + \varepsilon)(1 - \lambda) + \frac{10\lambda}{\ln t}, \quad \lambda = \frac{5}{2} \frac{1}{\ln^2 t} + O\left(\frac{1}{\ln^2 t}\right).$$

But we can now let $\varepsilon \to 0$, and we arrive at the contradictory inequality, for sufficiently large t and $\theta > 0$:

$$\theta \leqslant \theta(1-\lambda) + \frac{10\lambda}{\ln t} < \theta\left(1 - \frac{\lambda}{2}\right), \quad \ln t > \frac{20}{\theta}\,(t > t_0).$$

Thus $\theta = 0$ and $M(x) = o(x)$.

LEMMA 4. *If* $(l, D) = 1$, $D > 1$, *the asymptotic formula* $M(l, x) = o(x)$ *holds.*

The proof of this lemma in no way differs from the proof of Lemma 3. We obtain from inequalities (3.4.9), on putting $h = \varphi(D)$, $U_l(x) = 1/x\,|M(l, x)|$:

$$U_l(x) \leqslant \frac{1}{h \ln x} \sum_{q=1}^{D} \chi_0(q) \sum_{n \leqslant x} \frac{1}{n}\, U_q\left(\frac{x}{n}\right) + O\left(\frac{\ln \ln x}{\ln x}\right), \qquad (3.5.11)$$

which holds for any l, $(l, D) = 1$. We again obtain:

$$\theta = \overline{\lim_{x \to \infty}} \max_{1 \leqslant q \leqslant D} [\chi(q)\, U_q(x)].$$

Now, for any ε, $\theta > \varepsilon > 0$, $x \geqslant x_0(\varepsilon)$, $U_l(x) < \theta + \varepsilon$, $\theta - \varepsilon < \max_{1 \leqslant q \leqslant D} \chi(q)\, U_q(x_k)$ for an infinite set of unboundedly increasing x_k.

In accordance with Lemma 2, we choose the numbers q_k from the interval $t^{2k} \leqslant q_k \leqslant t^{2k+1}$, such that

$$|M_l(q_k)| < \frac{Cq_k}{\ln t}\,; \quad |M_l(q_k + v)| < \frac{Cq_k}{\ln t} + v < \frac{2Cq_k}{\ln t},$$

if $v \ln t < Cq_k$. Hence, for $0 \leqslant v \leqslant [Cq_k/\ln t]$:

$$U_l(q_k + v) < \frac{2C}{\ln t}.$$

But, in order for x/n to be situated within the limits

$$q_k \leqslant \frac{x}{n} \leqslant \left(1 + \frac{C}{\ln t}\right) q_k, \qquad (3.5.12)$$

the number n must satisfy

$$p_k = \left[\frac{1}{1 + \dfrac{C}{\ln t}}\,\frac{x}{q_k}\right] \leqslant n \leqslant \left[\frac{x}{q_k}\right] = p_k' < p_{k+1}. \qquad (3.5.13)$$

We have the expression, analogous to (3.5.9):

$$S_k = \sum_{p_k \leqslant n \leqslant p_k'} \frac{1}{n} = \ln \frac{p_k'}{p_k} + \alpha_k \frac{q_k}{x} = \ln\left(1 + \frac{C}{\ln t}\right) + \beta_k \frac{t^{2k+1}}{x}, \qquad (3.5.14)$$

where $\beta_k = O(1)$ with respect to x and t. Since, with n satisfying (3.5.13), we have $U_l(x) < 2C/(\ln t)$, and with

$$m = \left[\frac{1}{2}\frac{\ln x}{\ln t} - \frac{1}{2}\right],$$

$$\sum_1^m S_k = m \ln\left(1 + \frac{C}{\ln t}\right) + \sum_{k=1}^m \beta_k \frac{t^{2k+1}}{x} = m \ln\left(1 + \frac{C}{\ln t}\right) + O(1)$$

$$= \lambda \ln x + O(1), \lambda = \frac{\ln\left(1 + \frac{C}{\ln t}\right)}{2 \ln t}, \tag{3.5.15}$$

on replacing $U_q(x/n)$ by $2C/(\ln t)$ in (3.5.11) if n satisfies (3.5.13), and by $\theta + \varepsilon$ otherwise, we obtain the inequality

$$\max_{\leqslant l \leqslant D} [\chi_0(l)\, U_l(x)] < \frac{1}{h \ln x}\left[(\theta + \varepsilon)\Big[\sum_{n \leqslant x} \frac{1}{n} - \sum_{k=1}^m S_k\Big]\right.$$

$$\left. + \frac{2C}{\ln t} \sum_{k=1}^m S_k\right] + O\left(\frac{\ln \ln x}{\ln x}\right) = \frac{1}{h}(\theta + \varepsilon)(1 - \lambda)$$

$$+ \frac{2C\lambda}{\ln t} + O\left(\frac{\ln \ln x}{\ln x}\right), \quad h = \varphi(D). \tag{3.5.16}$$

On passing to the upper limit on the left, we obtain

$$\theta \leqslant \frac{1}{h}(\theta + \varepsilon)(1 - \lambda) + \frac{2C\lambda}{\ln t}$$

or, since $\varepsilon > 0$ is arbitrary,

$$\theta \leqslant \frac{1}{h}\theta(1 - \lambda) + \frac{2C\lambda}{\ln t} \leqslant \frac{1}{h}\theta\left(1 - \frac{\lambda}{2}\right),$$

which holds with $\ln t > 4Ch/\theta$.

This last inequality shows that $\theta = 0$, in other words, it proves our lemma.

Theorem 3.5.1 is an immediate consequence of Lemmas 1, 3 and 4.

THEOREM 3.5.1. *The asymptotic formulae*

$$\pi(x) = \frac{x}{\ln x} + o\left(\frac{x}{\ln x}\right) \tag{3.5.17}$$

and

$$\pi(l, D; x) = \frac{1}{\varphi(D)} \frac{x}{\ln x} + o\left(\frac{x}{\ln x}\right), \tag{3.5.18}$$

hold, where $(D, l) = 1$ *and* $\varphi(D)$ *is Euler's function.*

It may be mentioned that, if rather more accurate asymptotic formulae are introduced into Lemmas 1–4, $o[x/(\ln x)]$ in Theorem 3.5.1 can be replaced at least by

$$O\left(\frac{x}{\ln x}\sqrt{\frac{\ln\ln x}{\ln x}}\right).$$

But it seems impossible to bring about a substantial improvement in the remainder term in (3.5.17) and (3.5.18) while remaining within the framework of Selberg's method, and it is only with very great difficulty that it can be reduced to $O[x/(\ln^{2-\varepsilon}x)]$ ($\varepsilon > 0$ is arbitrary). The above method of proof of the limit theorems on the distribution of primes is due to Selberg,[46] while the idea of using $M(x)$ instead of $\psi(x) - x$ in Selberg's scheme for the proof of the law of distribution of primes in the natural series is due to Postnikov and Romanov.[22]

The elementary proof of Dirichlet's theorem is due to Shapiro.[47] In order to obtain in (3.5.17) and (3.5.18) a remainder term of the Hadamard–de la Vallée Poussin type, in other words, a remainder term of the form $0(x\,e^{-\lambda\ln^{\gamma}x})$, where $\lambda > 0$, $1 > \gamma > 0$ are constants, it is possible to utilize an asymptotic formula obtainable by elementary means (see [7]), namely,

$$\sum_{n\leqslant x}\frac{\chi(n)\,\mu(n)}{n^{1+i\tau}} = O[\ln^{7}(2+|\tau|)],$$

where χ is any character of the modulus D and τ is arbitrary; but unfortunately, the passage from this formula to the remainder term in (3.5.17) and (3.5.18) is not sufficiently elementary, since it requires contour integration. Consequently, instead of dwelling further on the remainder term, we shall occupy ourselves with some other variants of the elementary ideas described above.

6. Primes in Sequences of a Rather More General Type than Progressions

We shall give an elementary proof of a theorem on the infinity of primes in series other than progressions.

Let $\tau \neq 0$ and be real. We put

$$\alpha(\tau) = 1 + \frac{1}{i\tau} + \frac{l}{i\tau}\sum_{2}^{\infty}\left\{\frac{1}{i\tau}[n^{-i\tau} - (n-1)^{-i\tau}] + n^{-1-i\tau}\right\}. \quad (3.6.1)$$

This series is convergent, since

$$\frac{1}{i\tau}[n^{-i\tau}-(n-1)^{-i\tau}]+n^{-1-i\tau}=O\left(\frac{1}{n^2}\right). \tag{3.6.2}$$

Hence it follows at once that

$$\sum_{n\leqslant x}\frac{1}{n^{1+i\tau}}=\frac{i}{\tau}x^{-i\tau}+\alpha(\tau)+O\left(\frac{1}{x}\right). \tag{3.6.3}$$

We have further:

$$\begin{aligned}\sum_{n\leqslant x}\frac{\ln n}{n^{1+i\tau}}&=\ln[x]\sum_{n\leqslant x}\frac{1}{n^{1+i\tau}}-\sum_{n\leqslant x}\ln\left(1+\frac{1}{n}\right)\sum_{k\leqslant n}\frac{1}{k^{1+i\tau}}\\&=\frac{i}{\tau}x^{-i\tau}\ln x+\alpha(\tau)\ln x-\sum_{n\leqslant x}\frac{1}{n}\left[\frac{i}{\tau}n^{-i\tau}+\alpha(\tau)\right]\\&\quad+O(1)=\frac{i}{\tau}x^{-i\tau}\ln x+O(1).\end{aligned} \tag{3.6.4}$$

We consider the simple identity:

$$\begin{aligned}\sum_{n\leqslant x}\frac{\ln n}{n^{1+i\tau}}&=\sum_{n\leqslant x}\frac{1}{n^{1+i\tau}}\sum_{d\mid n}\Lambda(d)=\sum_{n\leqslant x}\frac{\Lambda(n)}{n^{1+i\tau}}\sum_{k\leqslant\frac{x}{n}}\frac{1}{k^{1+i\tau}}\\&=\frac{i}{\tau}\sum_{k\leqslant x}\frac{\Lambda(k)}{k^{1+i\tau}}\left(\frac{x}{k}\right)^{-i\tau}+\alpha(\tau)\sum_{n\leqslant x}\frac{\Lambda(n)}{n^{1+i\tau}}+\sum_{k\leqslant x}\frac{\Lambda(k)}{k^{1+i\tau}}O\left(\frac{k}{x}\right)\\&=\frac{i}{\tau}x^{-i\tau}\ln x+\alpha(\tau)\sum_{n\leqslant x}\frac{\Lambda(n)}{n^{1+i\tau}}+O(1),\end{aligned} \tag{3.6.5}$$

which is obtained by making use of (3.1.6) and (3.1.7).

It will be seen, on combining (3.5.4) and (3.6.5), that

$$\alpha(\tau)\sum_{n\leqslant x}\frac{\Lambda(n)}{n^{1+i\tau}}=O(1). \tag{3.6.6}$$

We shall consider a further identity, namely:

$$\begin{aligned}\sum_{n\leqslant x}\frac{1}{n^{1+i\tau}}\sum_{d\mid n}\mu(d)\ln\frac{x}{d}&=\ln x+\sum_{n\leqslant x}\frac{\Lambda(n)}{n^{1+i\tau}}\\&=\sum_{n\leqslant x}\frac{\mu(n)}{n^{1+i\tau}}\ln\frac{x}{n}\sum_{k\leqslant\frac{x}{n}}\frac{1}{k^{1+i\tau}}=\frac{i}{\tau}x^{-i\tau}\sum_{n\leqslant x}\frac{\mu(n)}{n}\ln\frac{x}{n}\\&\quad+\alpha(\tau)\sum_{n\leqslant x}\frac{\mu(n)}{n^{1+i\tau}}\ln\frac{x}{n}+\sum_{n\leqslant x}\frac{\mu(n)}{n^{1+i\tau}}\ln\frac{x}{n}O\left(\frac{n}{x}\right)\\&=\alpha(\tau)\sum_{n\leqslant x}\frac{\mu(n)}{n^{1+i\tau}}\ln\frac{x}{n}+O(1),\end{aligned} \tag{3.6.7}$$

since

$$\frac{1}{x}\sum_{n\leqslant x}\ln\frac{x}{n}=O(1) \tag{3.6.8}$$

and

$$\sum_{n\leqslant x}\frac{\mu(n)}{n}\ln\frac{x}{n}=\frac{1}{x}\sum_{n\leqslant x}\mu(n)\left[\frac{x}{n}\right]\ln\frac{x}{n}-\frac{1}{x}\sum_{n\leqslant x}\mu(n)\left\{\frac{x}{n}\right\}\ln\frac{x}{n}$$

$$=\frac{1}{x}\sum_{n\leqslant x}\sum_{d\mid n}\mu(d)\ln\frac{x}{d}+O(1)$$

$$=\frac{1}{x}\sum_{n\leqslant x}\Lambda(n)+O(1)=O(1). \tag{3.6.9}$$

On combining (3.6.6) and (3.6.7), we see that

$$\varphi(\tau,x)=\sum_{n\leqslant x}\frac{\Lambda(n)}{n^{1+i\tau}}=\delta(\tau)\ln x+O(1), \tag{3.6.10}$$

where

$$\delta(\tau)=\begin{cases}0, & \alpha(\tau)\neq 0,\\ -1, & \alpha(\tau)=0.\end{cases} \tag{3.6.11}$$

Let us show that, given any $\tau\neq 0$, $\delta(\tau)=0$, in other words, we always have $\alpha(\tau)\neq 0$. In fact,

$$\operatorname{Re}\left[3\varphi(0,x)+4\varphi(\tau,x)+\varphi(2\tau,x)\right]=$$

$$=2\sum_{n\leqslant x}\frac{\Lambda(n)}{n^{1+i\tau}}\left[\cos(\tau\ln n)+1\right]^2$$

$$=\left[3+4\delta(\tau)+\delta(2\tau)\right]\ln x+O(1), \tag{3.6.12}$$

whence it follows that $\delta(\tau)=-1$, since the left-hand side is non-negative. Hence $\delta(\tau)=0$, $\tau\neq 0$ and $\alpha(\tau)\neq 0$.

Thus, if $\tau\neq 0$, we have

$$\sum_{n\leqslant x}\frac{\Lambda(n)}{n^{1+i\tau}}=O(1),\ \tau\neq 0;\quad \sum_{n\leqslant x}\frac{\Lambda(n)}{n}=\ln x+O(1). \tag{3.6.13}$$

Now let a and α be real, $a>0$, and let $P(x)$ be a polynomial:

$$P(z)=\sum_{k=-s}^{k=s}a_k z^k,\qquad a_0=\int_0^1 P(e^{2\pi it})\,dt.$$

We have, by virtue of (3.6.13):

$$\sum_{n\leqslant x}P\left[\left(\frac{n}{a}\right)^{2\pi i\alpha}\right]\frac{\Lambda(n)}{n}=\sum_{k=-s}^{s}a_k a^{-2\pi ik\alpha}\sum_{n\leqslant x}\frac{\Lambda(n)}{n^{1-2\pi ik\alpha}}$$

$$=a_0\ln x+O(1). \tag{3.6.14}$$

Let $U(z)$ be a periodic real function of period 1, $U(z+1) = U(z)$, and let it be continuous in $[0, 1]$ except for a finite number of discontinuities of the first kind. It is well known that, given such a function, and any $\varepsilon > 0$, two polynomials $P_1(z)$, $P_2(z)$ can be found such that

$$P_1(e^{2\pi iz}) \leqslant U(z) \leqslant P_2(e^{2\pi iz}), \qquad \int_0^1 |P_1(e^{2\pi iz}) - P_2(e^{2\pi iz})| dz < \varepsilon.$$

Hence

$$\sum_{n\leqslant x} P_1\left[\left(\frac{n}{a}\right)^{2\pi i\alpha}\right]\frac{\Lambda(n)}{n} \leqslant \sum_{n\leqslant x} U\left[\alpha \ln\frac{n}{a}\right]\frac{\Lambda(n)}{n} \leqslant \sum_{n\leqslant x} P_2\left[\left(\frac{n}{a}\right)^{2\pi i\alpha}\right]\frac{\Lambda(n)}{n},$$

from which it follows, in view of (3.6.14), that

$$\int_0^1 P_1(e^{2\pi it})\, dt \ln x + O(1) \leqslant \sum_{n\leqslant x} U\left(\alpha \ln\frac{n}{a}\right)\frac{\Lambda(n)}{n} \leqslant \int_0^1 P_2(e^{2\pi it})\, dt \ln x + O(1).$$

Consequently, since ε can be as small as desired, we can assert that

$$\sum_{n\leqslant x} U\left[\alpha \ln\frac{n}{a}\right]\frac{\Lambda(n)}{n} = \sum_{p\leqslant x} U\left(\alpha \ln\frac{p}{a}\right)\frac{\ln p}{p} + O(1)$$

$$= \int_0^1 U(t)\, dt \ln x + O(\ln x). \tag{3.6.15}$$

Relationship (3.6.15) can be used for proving a theorem of the Dirichlet type for a class of numerical series which was earlier discussed by Gel'fond†.

Let $m > 1$, $r \geqslant 1$, $q \geqslant 1$ be integers. We consider the group of substitutions

$$L_k(x) = mx + qm + k \qquad (k = 0, 1, \ldots, m-1).$$

We substitute r for x in our linear forms. We obtain m numbers of the first group. These m numbers are now substituted in our

† See [7] in the References.

linear forms. We obtain m^2 integers. We continue this process indefinitely and obtain a numerical series L. All the numbers of this series are easily seen to be different: this follows at once from the form of any number N belonging to the series L, namely

$$N = m^s r + qm\frac{m^s - 1}{m - 1} + \sum_{i=0}^{s-1} k_i m^i \qquad (s = 1,2\cdots),$$

where k_i $(i = 0, 1, \ldots)$ are arbitrary integers, $0 \leqslant k_i \leqslant m - 1$. Hence N belongs to the series L if

$$m^s r + qm\frac{m^s - 1}{m - 1} \leqslant N \leqslant m^s(r + 1) + qm\frac{m^s - 1}{m - 1} - 1 \qquad (3.6.16)$$

with any $s \geqslant 1$.

Inequalities (3.6.16) are necessary and sufficient conditions for the number N to belong to the series L. The properties of series of the type L are considered in detail in [7]. It may readily be seen that the set of integers satisfying the conditions

$$N = m^s\left(r + \frac{qm}{m - 1}\right) + t \qquad (0 \leqslant t \leqslant m^s, s = 1, 2, \ldots), \quad (3.6.17)$$

where t is an integer, only differs from the set of numbers defined by inequalities (3.6.16) by a set of the density of a geometric progression. But conditions (3.6.17) can be rewritten as

$$\frac{\ln \dfrac{N}{r + \dfrac{qm}{m - 1}}}{\ln m} = s + \frac{\ln\left[1 + \dfrac{t}{\left(r + \dfrac{qm}{m - 1}\right) m^s}\right]}{\ln m} \quad (0 \leqslant t \leqslant m^s),$$

in other words, as

$$\left\{\frac{\ln \dfrac{N}{r + \dfrac{qm}{m - 1}}}{\ln m}\right\} \leqslant \frac{1}{\ln m}\left\{\ln\left(1 + \frac{1}{r + \dfrac{qm}{m - 1}}\right)\right\}. \qquad (3.6.18)$$

We now pose the question of the number of primes in the series L. Since the condition that a number N belong to L, discounting a set of numbers forming a geometric progression, is (3.6.18), we

define a function $U(x)$ by putting

$$U(x) = \begin{cases} 1, & 0 \leqslant x \leqslant \dfrac{1}{\ln m} \ln\left(1 + \dfrac{1}{r + \dfrac{qm}{m-1}}\right) = x_0, \\ 0, & x_0 < x < 1, \end{cases}$$

$$U(x+1) = U(x).$$

The condition for N to belong to L, discounting a very sparse set, is now of the form

$$U\left[\frac{\ln\left(\dfrac{N}{r + \dfrac{qm}{m-1}}\right)}{\ln m}\right] = 1.$$

On making use of (3.3.8), we obtain the theorem:

$$\sum_{\substack{p<x \\ p\in L}} \frac{\ln p}{p} = \frac{1}{\ln m} \ln\left(1 + \frac{1}{r + \dfrac{qm}{m-1}}\right) \ln x + o(\ln x),$$

a particular consequence of which is that the series L contains an infinity of primes. It follows naturally that, if we use a law of distribution of primes with a good remainder term, we can find the number of primes as far as x in the series L to a high accuracy. Formula (3.6.16) can also be proved by elementary means for an arithmetic progression, without any added complexities. We have in this case the relationship

$$\sum_{\substack{p<x \\ p\equiv l \pmod D}} \frac{\ln p}{p} U\left[\alpha \ln \frac{n}{a}\right] = \frac{1}{\varphi(D)} \int_0^1 U(t)\, dt \ln x + o(\ln x),$$

whence the formula

$$\sum_{\substack{p<x \\ p\equiv l \pmod D,\, p\in L}} \frac{\ln p}{p} = \frac{1}{\varphi(D) \ln m} \ln\left(1 + \frac{1}{r + \dfrac{qm}{m-1}}\right) \ln x + o(\ln x) \tag{3.6.19}$$

also follows.

We prove a generalization of (3.6.13), from which (3.6.19) will follow immediately. In fact, if $\chi \neq \chi_0$ and $\tau \neq 0$, it is obvious that

$$\sum_{n<x} \frac{\chi(n)}{n^{1+i\tau}} = \alpha + O\left(\frac{1}{x}\right), \quad \alpha = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^{1+i\tau}}. \tag{3.6.20}$$

Furthermore, we have

$$\sum_{n\leqslant x}\frac{\chi(n)\ln n}{n^{1+i\tau}}=\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n^{1+i\tau}}\sum_{k\leqslant x/n}\frac{\chi(k)}{k^{1+i\tau}}$$

$$=\alpha\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n^{1+i\tau}}+O(1)=O(1) \tag{3.6.21}$$

in view of the convergence of the series on the left-hand side of (3.6.21).

Similarly, we obtain

$$\sum_{n\leqslant x}\frac{\chi(n)}{n^{1+i\tau}}\sum_{d\mid n}\mu(d)\ln\frac{x}{d}=\ln x+\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n^{1+i\tau}}$$

$$=\sum_{n\leqslant x}\frac{\chi(n)\mu(n)}{n^{1+i\tau}}\ln\frac{x}{n}\sum_{k\leqslant x/n}\frac{\chi(k)}{k^{1+i\tau}}$$

$$=\alpha\sum_{n\leqslant x}\frac{\chi(n)\mu(n)}{n^{1+i\tau}}\ln\frac{x}{n}+O(1) \tag{3.6.22}$$

by virtue of (3.6.20) and the fact that $\sum_{n\leqslant x}\ln x/n=O(x)$. On combining (3.6.21) and (3.6.22), it will be seen that we always have, with $\tau\neq 0$ and $\chi\neq\chi_0$:

$$\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n^{1+i\tau}}=\delta(\chi,\tau)\ln x+O(1) \tag{3.6.23}$$

$$\delta(\chi,\tau)=\begin{cases}0, & \alpha\neq 0,\\ -1, & \alpha=0.\end{cases}$$

We observe also that, with $\tau\neq 0$,

$$\sum_{n\leqslant x}\frac{\chi_0(n)\Lambda(n)}{n^{1+i\tau}}=\sum_{n\leqslant x}\frac{\Lambda(n)}{n^{1+i\tau}}+O(1)=O(1)$$

by virtue of (3.6.13). Let us show that $\delta(\chi,\tau)=0$ always. We have, first of all,

$$\operatorname{Re}\left[3\sum_{n\leqslant x}\frac{\chi_0(n)\Lambda(n)}{n}+4\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n^{1+i\tau}}+\sum_{n\leqslant x}\frac{\chi^2(n)\Lambda(n)}{n^{1+2i\tau}}\right]$$

$$=2\sum_{n\leqslant x}\frac{\chi(n)\Lambda(n)}{n}\left[1+\operatorname{Re}\frac{\chi(n)}{n^{i\tau}}\right]^2\geqslant 0, \tag{3.6.24}$$

whence it follows that

$$[3+4\delta(\chi,\tau)+\delta(\chi^2,2\tau)]\ln x+O(1)\geqslant 0.$$

Thus, since $\delta(\chi^2, 2\tau) = -1, 0$, then $\delta(\chi, \tau) \neq -1$, since otherwise the left-hand side of the last inequality would be negative for large x. We therefore always have

$$\sum_{n \leqslant x} \frac{\chi(n)\Lambda(n)}{n^{1+i\tau}} = O(1), \tag{3.6.25}$$

provided either $\chi \neq \chi_0$ or $\tau \neq 0$. Formula (3.6.19) now follows at once from this.

CHAPTER 4

AN ELEMENTARY DERIVATION OF THE LAW OF DISTRIBUTION OF GAUSSIAN PRIMES. A THEOREM ON ALMOST—PRIME GAUSSIAN NUMBERS

1. Introduction

The present chapter assumes a knowledge of the elements of the arithmetic of complex ("Gaussian") integers, of the form $a + bi$, where $i^2 = -1$; a, b are ordinary integers. (In this connection, see, for example, the small book by Schnirelmann.[31]) Such integers include in particular the Gaussian primes, such as $1 + i$, 3, $2 \pm i$, etc. Our interest is in asymptotic formulae for the number of Gaussian primes inside an expanding contour.

For certain contours, the problem reduces to the law, considered above by elementary methods, of the distribution of primes in an arithmetic progression (see Chapter 3). For instance, the number of Gaussian primes ϱ inside a circle $|z| < R$ is asymptotically equal to eight times the number of primes of the form $x^2 + y^2 < R^2$. These primes coincide with the number 2 or with primes of the form $4n + 1$. The number of such numbers not exceeding R^2 is given asymptotically by

$$\frac{R^2}{\varphi(4) \ln R^2} = \frac{R^2}{4 \ln R},$$

so that the number required by us is asymptotically equal to

$$\frac{2R^2}{\ln R}.$$

In the case of a contour of general type, however, the problem does not reduce to the distribution of ordinary primes in a pro-

gression. The solution of this problem to be described below was given in 1956 by Chulanovskii,[25] using Selberg's method.

We shall adopt the following notation. Small Greek letters (except π) will denote Gaussian integers, the letter ξ any complex number (not necessarily an integer), the letters ϱ, σ, τ Gaussian primes. The symbol $[\xi]$ will denote a Gaussian integer such that

$$\mathrm{Re}[\xi] - \tfrac{1}{2} < \mathrm{Re}\xi \leqslant \mathrm{Re}[\xi] + \tfrac{1}{2}$$

and

$$\mathrm{Im}[\xi] - \tfrac{1}{2} < \mathrm{Im}\xi \leqslant \mathrm{Im}[\xi] + \tfrac{1}{2}.$$

D is a domain on the plane of the complex variable, $R = \max_{\xi \in D} |\xi|$, M_L is the set of Gaussian integers λ such that there exists on the line L a ξ such that $[\xi] = \lambda$; $z_L = \sum_{\lambda \in M_L} 1$. Finally, we denote by the sign D/ν the domain defined as follows: $\xi \in D/\nu$ when and only when $\xi\nu \in D$; similarly, L/ν denotes the line such that $\xi \in (L/\nu)$ when and only when $\xi\nu \in L$.

We shall denote real constants by $c_1, c_2, c_3, \ldots$

We shall give an elementary proof of the following theorem.

THEOREM 4.1.1 (Chulanovskii). *If there exists a line L, which is connected (or consists of $O(1)$ connected pieces), contains the boundary of the domain D, and is such that $Z_L = BR$, then*

$$\sum_{\varrho \in D} 1 = \frac{2}{\pi \ln R} \sum_{\nu \in D} 1 + O\left(\frac{R^2}{\ln R \sqrt[3]{\ln \ln R}}\right). \tag{4.1.1}$$

We can obviously assume without loss of generality that the domain D is symmetric with respect to rotation through 90° (i.e. is such that $\xi \in D$ implies $i\xi \in D$, $-\xi \in D$, $-i\xi \in D$). We shall assume such a domain in order to facilitate our arguments.

It should be noted that, by using the theory of analytic functions (specifically, by using Hecke series), a distribution law can be obtained with a much better remainder term (see, for example, Kubilyus[9]. Admittedly, Kubilyus proved his theorem for homothetically expanding contours, but it is not difficult to see that this restriction is not essential and can easily be lifted). The proof requires the use of a denumerable number of Hecke series. The value of the present method of Selberg lies in the fact that it is completely elementary.

We shall require a function, denoted by $q(\alpha)$ and defined as follows (for $\alpha \neq 0$):

$$q(\alpha) = \begin{cases} 1, \text{ if } \alpha \text{ is one of the unities } (1, i, -1, -i), \\ 0, \text{ if } \alpha \text{ is divisible by the square of a number with modulus} > 1, \\ (-1)^k, \text{ if } \alpha = \varrho_1 \dots \varrho_k \text{ and the primes } \varrho_1, \dots, \varrho_k \text{ are non-associated.} \end{cases} \tag{4.1.2}$$

This function plays a role for Gaussian integers similar to the role of the Möbius function for rational integers, and has similar properties.

2. Some Auxiliary Formulae

During the proof of the theorem we shall require certain auxiliary formulae, which it is more convenient to derive first.

We start by noting some properties of the function $q(\alpha)$. First, it is obvious that

$$\sum_{\alpha|\varkappa} q(\alpha) = \begin{cases} 4, & \text{if } \varkappa \text{ is one of the unities} \\ 0, & \text{if } |\varkappa| > 1. \end{cases} \tag{4.2.1}$$

Further,

$$\sum_{1\leqslant|\alpha|\leqslant r} \frac{q(\alpha)}{|\alpha|^2} = \frac{1}{\pi r^2} \sum_{1\leqslant|\varkappa|\leqslant r} \left(\sum_{1\leqslant\beta\leqslant r/|\alpha|} 1 + O\left(\frac{r}{|\alpha|}\right)\right) q(\alpha)$$

$$= \frac{1}{\pi r^2} \sum_{1\leqslant|\nu|\leqslant r} \sum_{\alpha|\nu} q(\alpha) + O\left(\frac{1}{r} \sum_{1\leqslant|\alpha|\leqslant r} \frac{1}{|\alpha|}\right) = O(1),$$

so that

$$\sum_{1\leqslant|\alpha|\leqslant r} \frac{q(\alpha)}{|\alpha|^2} = O(1). \tag{4.2.2}$$

It is also obvious that

$$\sum_{1\leqslant|\alpha|\leqslant r} q(\alpha) = O(r^2). \tag{4.2.3}$$

Certain formulae concerning primes will also be required. First, it is obvious that

$$\ln|\nu| = \sum_{\substack{\varrho,k \\ \varrho^k|\nu}} \frac{\ln|\varrho|}{4}. \tag{4.2.4}$$

Also, since

$$\sum_{\substack{\varrho,k \\ |\varrho^k|\leqslant r}} \ln|\varrho| = O(r^2) \tag{4.2.5}$$

(this is a simple consequence of the familiar connection between Gaussian primes and rational primes, but could be proved without

a reference to rational primes), it is easily seen that

$$\sum_{\substack{\varrho,k \\ |\varrho^k|<r}} \frac{\ln|\varrho|}{|\varrho^k|} = O(r). \tag{4.2.6}$$

Hence

$$\begin{aligned}\sum_{1<|\nu|<r} \ln|\nu| &= \sum_{1<|\nu|<r} \sum_{\substack{\varrho,k \\ \varrho^k|\nu}} \frac{\ln|\varrho|}{4} \\ &= \frac{1}{4} \sum_{\substack{\varrho,k \\ |\varrho^k|<r}} \ln|\varrho| \sum_{\substack{1<|\nu|<r \\ \nu:\varrho^k}} 1 = \frac{1}{4} \sum_{\substack{\varrho,k \\ |\varrho^k|<r}} \ln|\varrho| \left(\frac{\pi r^2}{|\varrho^k|^2} + O\left(\frac{r}{|\varrho^k|}\right)\right) \\ &= \frac{\pi r^2}{4} \sum_{\substack{\varrho,k \\ |\varrho^k|<r}} \frac{\ln|\varrho|}{|\varrho^k|^2} + O\left(r \sum_{\substack{\varrho,k \\ |\varrho^k|<r}} \frac{\ln|\varrho|}{|\varrho^k|}\right) \\ &= \frac{\pi r^2}{4} \sum_{\substack{\varrho,k \\ |\varrho^k|<r}} \frac{\ln|\varrho|}{|\varrho^k|^2} + O(r^2).\end{aligned}$$

It follows from this, and the fact that

$$\sum_{1<|\nu|<r} \ln|\nu| = \pi r^2 \ln r + O(r^2)$$

that

$$\sum_{\substack{\varrho,k \\ |\varrho^k|<r}} \frac{\ln|\varrho|}{|\varrho^k|^2} = 4\ln r + O(1). \tag{4.2.7}$$

Since

$$\sum_{\substack{\varrho,k \\ |\varrho^k|<r,\, k>2}} \frac{\ln|\varrho|}{|\varrho^k|^2} = O(1),$$

we also have

$$\sum_{|\varrho|<r} \frac{\ln|\varrho|}{|\varrho|^2} = 4\ln r + O(1). \tag{4.2.8}$$

We shall also require the formulae

$$\sum_{1<|\nu|<r} \frac{\ln|\nu|}{|\nu|^2} = \varrho \ln^2 r + c_1 + O\left(\frac{\ln r}{r}\right), \tag{4.2.9}$$

$$\sum_{1<|\nu|<r} \frac{1}{|\nu|^2} = 2\pi \ln r + c_2 + O\left(\frac{1}{r}\right). \tag{4.2.10}$$

Finally, we shall prove the asymptotic formula

$$\sum_{\substack{\mu \\ \text{ex. } \lambda \in M_L \\ |\mu-\lambda|<r}} 1 = Br(Z_L + r), \tag{4.2.11}$$

which holds for $r \geqslant 1$. It is essential here that the line L be connected or consist of $O(1)$ connected pieces. We shall prove (4.2.11) on the assumption that L is connected.

The proof is extremely simple. If μ is such that there exists $\lambda \in M_L$ under the condition $|\mu - \lambda| \leqslant r$, then since L is connected, there exist numbers $\lambda' \in M_L \gg \min(Z_L, r)$, such that $|\mu - \lambda'| \leqslant 2r$. Hence

$$\sum_{\substack{\mu \\ \text{ex. } \lambda \in M_L \\ |\mu-\lambda| \leqslant r}} \sum_{\substack{\lambda' \in M_L \\ |\mu-\lambda'| \leqslant 2r}} 1 \gg \min(Z_L, r) \sum_{\substack{\mu \\ \text{ex. } \lambda \in M_L \\ |\mu-\lambda| \leqslant r}} 1.$$

On the other hand,

$$\sum_{\substack{\mu \\ \text{ex. } \lambda \in M_L \\ |\mu-\lambda| \leqslant r}} \sum_{\substack{\lambda' \in M_L \\ |\mu-\lambda'| \leqslant 2r}} 1 \leqslant \sum_{\lambda' \in M_L} \sum_{\substack{\mu \\ |\mu-\lambda'| \leqslant 2r}} 1 \ll Z_L r^2.$$

Hence

$$\sum_{\substack{\mu \\ \text{ex. } \lambda \in M_L \\ |\mu-\lambda| \leqslant r}} 1 \ll \frac{Z_L r^2}{\min(Z_L, r)} \ll r(Z_L + r),$$

and (4.2.11) is proved. The proof follows the same lines when L consists of $O(1)$ connected pieces. We carry out the argument for each piece separately then add.

It obviously follows from (4.2.11) (with $\nu \neq 0$) that

$$\sum_{\mu \in D/\nu} 1 = \frac{1}{|\nu|^2} \sum_{\varkappa \in D} 1 + O\left(\frac{Z_L}{|\nu|} + 1\right), \tag{4.2.12}$$

as also, that

$$\sum_{\lambda \in M_L/\nu} 1 \ll \frac{Z_L}{\nu} + 1. \tag{4.2.13}$$

It follows from (4.2.13) and from the fact that $Z_L \ll R$, that, when $|\nu| \leqslant R$,

$$\sum_{\lambda \in M_L/\nu} 1 \ll \frac{R}{|\nu|}. \tag{4.2.14}$$

The condition imposed on D is therefore also fulfilled for D/ν so that we are justified in applying to D/ν the formulae to be obtained for D. We shall make use of this fact later.

3. Proof of the Formula for

$$\sum_{\varrho\in D} \ln^2|\varrho| + \tfrac{1}{4}\sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in D}} \ln|\varrho|\ln|\sigma|$$

We shall prove the following lemma in this section.

LEMMA 1. *The following formula holds*:

$$\sum_{\varrho\in D} \ln^2|\varrho| + \frac{1}{4}\sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in D}} \ln|\varrho|\ln|\sigma| = \frac{4\ln R}{\pi}\sum_{\varkappa\in D} 1 + O(R^2), \quad (4.3.1)$$

which can be written alternatively as

$$\sum_{\varrho\in D} \ln^2|\varrho| + \frac{1}{2}\sum_{|\varrho|\leqslant\sqrt{R}} \ln|\varrho| \sum_{\sigma\in D/\varrho} \ln|\sigma| = \frac{4\ln R}{\pi}\sum_{\varkappa\in D} 1 + O(R^2). \quad (4.3.2)$$

The equivalence of (4.3.1) and (4.3.2) is clear from the fact that

$$\sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in D}} \ln|\varrho|\ln|\sigma| = \sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in D\\ |\varrho|\leqslant\sqrt{R}}} \ln|\varrho|\ln|\sigma| + \sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in D\\ |\sigma|\leqslant\sqrt{R}}} \ln|\varrho|\ln|\sigma| - \sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in D\\ |\varrho|\leqslant\sqrt{R},\ |\sigma|\leqslant\sqrt{R}}} \ln|\varrho|\ln|\sigma| = 2\sum_{|\varrho|\leqslant\sqrt{R}} \ln|\varrho| \sum_{\sigma\in D/\varrho} \ln|\sigma| + O(R^2)$$

by virtue of (4.3.2).

For the proof of Lemma 1, we subject the sum

$$\sum_{\substack{\nu\in D/\alpha\\ \nu\neq 0}} \ln^2|\nu|, \quad \text{where} \quad \alpha\neq 0,$$

to certain transformations.

On the one hand, by (4.2.4),

$$\sum_{\substack{\nu\in D/\alpha\\ \nu\neq 0}} \ln^2|\nu| = \sum_{\substack{\nu\in D/\alpha\\ \nu\neq 0}} \sum_{\varrho,\sigma,k,m\atop \varrho^k|\nu,\sigma^m|\nu} \frac{\ln|\sigma|\ln|\sigma|}{16}$$

$$= \sum_{\varrho,\sigma,k,m} \frac{\ln|\varrho|\ln|\sigma|}{16} \sum_{\substack{\nu\in D/\alpha\\ \nu\neq 0\\ \nu:\{\varrho^k,\sigma^m\}}} 1 = \sum_{\varrho,\sigma,k,m} \frac{\ln|\varrho|\ln|\sigma|}{16} \sum_{\substack{\mu\in D/\alpha\{\varrho^k,\sigma^m\}\\ \mu\neq 0}} 1$$

(where $\{\varrho^k, \sigma^m\}$ denotes the least common multiple of the numbers ϱ^k, σ^m, and where, in view of the symmetry of D with respect to rotation through 90°, it is a matter of indifference which of the

four common multiples is understood). Hence, in view of (4.2.1),

$$\sum_{\substack{|\alpha|\leqslant R\\ \alpha\neq 0}} q(\alpha) \sum_{\substack{\nu\in D/\alpha\\ \nu\neq 0}} \ln^2|\nu| = \sum_{\substack{|\alpha|\leqslant R\\ \alpha\neq 0}} q(\alpha) \sum_{\varrho,\sigma,k,m} \frac{\ln|\varrho|\ln|\sigma|}{16} \sum_{\substack{\mu\in D/\alpha\{\varrho^k,\sigma^m\}\\ \mu\neq 0}} 1$$

$$= \sum_{\varrho,\sigma,k,m} \frac{\ln|\varrho|\ln|\sigma|}{16} \sum_{\substack{\varkappa\in D/\{\varrho^k,\sigma^m\}\\ \varkappa\neq 0}} \sum_{\alpha|\varkappa} q(\alpha) = \sum_{\substack{\varrho,\sigma,k,m\\ \{\varrho^k,\sigma^m\}\in D}} \ln|\varrho|\ln|\sigma|.$$

On the other hand, on using (4.2.11) and (4.2.9), we find that

$$\sum_{\substack{\nu\in D/\alpha\\ \nu\neq 0}} \ln^2|\nu| = \frac{1}{|\alpha|^2} \sum_{\substack{\varkappa\\ \alpha[\varkappa/\alpha]\in D\\ [\varkappa/\alpha]\neq 0}} \ln^2\left|\left[\frac{\varkappa}{\alpha}\right]\right| = \frac{1}{|\alpha^2|} \sum_{\substack{\varkappa\in D\\ [\varkappa\frown\alpha]\neq 0}} \ln^2\frac{\varkappa}{\alpha}$$

$$+ \frac{1}{|\alpha|^2} \sum_{\substack{\varkappa\in D\\ [\varkappa/\alpha]\neq 0}} \left(\ln^2\left|\left[\frac{\varkappa}{\alpha}\right]\right| - \ln^2\left|\frac{\varkappa}{\alpha}\right|\right)$$

$$+ \frac{1}{|\alpha|^2}\left(\sum_{\substack{\varkappa\\ \alpha[\varkappa/\alpha]\in D\\ [\varkappa/\alpha]\neq 0}} \ln^2\left|\left[\frac{\varkappa}{\alpha}\right]\right| - \sum_{\substack{\varkappa\in D\\ [\varkappa/\alpha]\neq 0}} \ln^2\left|\left[\frac{\varkappa}{\alpha}\right]\right|\right)$$

$$= \frac{1}{|\alpha|_2} \sum_{\substack{\varkappa\in D\\ [\alpha/\varkappa]\neq 0}} \left(\frac{1}{\pi} \sum_{\substack{|\nu|\leqslant|\varkappa/\sigma|\\ \nu\neq 0}} \frac{\ln|\nu|}{|\nu|^2} - \frac{c_1}{\pi} + O\left(\frac{|\alpha|}{|\varkappa|}\ln\left(1+\frac{|\varkappa|}{|\alpha|}\right)\right)\right)$$

$$+ O\left(\frac{1}{|\alpha|^2} \sum_{\substack{\varkappa\in D\\ [\varkappa/\alpha]\neq 0}} \frac{|\alpha|}{|\varkappa|}\ln\left(1+\frac{|\varkappa|}{|\alpha|}\right)\right) + O\left(\frac{|\alpha|(Z_L+|\alpha|)}{|\alpha|^2}\ln^2\frac{2R}{|\alpha|}\right)$$

$$= \frac{1}{\pi|\alpha|^2} \sum_{\substack{\varkappa\in D\\ \varkappa\neq 0}} \sum_{\substack{|\nu|\leqslant|\varkappa/\alpha|\\ \nu\neq 0}} \frac{\ln|\nu|}{|\nu|^2} - \frac{c_1}{\pi|\alpha|^2} \sum_{\substack{\varkappa\in D\\ [\varkappa/\alpha]\neq 0}} 1$$

$$+ O\left(\sum_{\substack{\varkappa\in D\\ [\varkappa/\alpha]\neq 0}} \frac{1}{|\varkappa\alpha|}\ln\left(1+\frac{|\varkappa|}{|\alpha|}\right)\right) + O\left(\left(\frac{Z_L}{|\alpha|}+1\right)\ln^2\frac{2R}{|\alpha|}\right).$$

We work out

$$\sum_{\substack{\alpha\\ |\alpha|\leqslant R\\ \alpha\neq 0}} \frac{q(\alpha)}{\pi|\alpha|_2} \sum_{\substack{\varkappa\in D\\ \varkappa\neq 0}} \sum_{\substack{|\nu|\leqslant|\varkappa/\alpha|\\ \nu\neq 0}} \frac{\ln|\nu|}{|\nu|^2}$$

$$= \frac{1}{\pi} \sum_{\substack{\varkappa\in D\\ \varkappa\neq 0}} \sum_{\substack{|\alpha|\leqslant|\varkappa|\\ \alpha\neq 0}} \frac{q(\alpha)}{|\alpha|^2} \sum_{\substack{|\nu|\leqslant|\varkappa/\alpha|\\ \nu\neq 0}} \frac{1}{|\nu|^2} \sum_{\substack{\varrho,k\\ \varrho^k|\nu}} \frac{\ln|\varrho|}{4}$$

$$= \frac{1}{\pi} \sum_{\substack{\varkappa\in D\\ \varkappa\neq 0}} \sum_{\substack{\varrho,k\\ |\varrho^k|\leqslant|\varkappa|}} \frac{\ln|\varrho|}{4} \sum_{\substack{|\alpha|\leqslant|\varkappa/\varrho^k|\\ \alpha\neq 0}} \frac{q(\alpha)}{|\alpha|^2} \sum_{\substack{|\nu|\leqslant|\varkappa/\alpha|\\ \nu\neq 0\\ \nu:\varrho^k}} \frac{1}{|\nu|^2}$$

$$= \frac{1}{\pi} \sum_{\substack{\varkappa \in D \\ \varkappa \neq 0}} \sum_{\substack{\varrho, k \\ |\varrho^k| \leqslant |\varkappa|}} \frac{\ln |\varrho|}{4|\varrho^k|^2} \sum_{\substack{|\alpha| \leqslant |\varkappa/\varrho^k| \\ \alpha \neq 0}} \frac{q(\alpha)}{|\alpha|^2} \sum_{\substack{|\mu| \leqslant |\varkappa|/|\alpha\varrho^k| \\ \mu \neq 0}} \frac{1}{|\mu|^2}$$

$$= \frac{1}{\pi} \sum_{\substack{\varkappa \in D \\ \varkappa \neq 0}} \sum_{\substack{\varrho, k \\ |\varrho^k| \leqslant |\varkappa|}} \frac{\ln |\varrho|}{4|\varrho^k|^2} \sum_{\substack{|\nu| \leqslant |\varkappa/\varrho^k| \\ \nu \neq 0}} \frac{1}{|\nu|^2} \sum_{\alpha | \nu} q(\alpha)$$

$$= \frac{1}{\pi} \sum_{\substack{\varkappa \in D \\ \varkappa \neq 0}} (16 \ln |\varkappa| + O(1)) = \frac{16 \ln R}{\pi} \sum_{\varkappa \in D} 1 + O(R)^2$$

by (4.2.4), (4.2.1) and (4.2.7);

$$\sum_{\substack{|\alpha| \leqslant R \\ \alpha \neq 0}} q(\alpha) \frac{c_1}{\pi |\alpha|^2} \sum_{\substack{\varkappa \in D \\ [\varkappa/\alpha] \neq 0}} 1$$

$$= \frac{c_1}{\pi} \sum_{\substack{|\alpha| \leqslant R \\ \alpha \neq 0}} \frac{q(\alpha)}{|\alpha|^2} \sum_{\varkappa \in D} 1 - \frac{c_1}{\pi} \sum_{\substack{|\alpha| \leqslant R \\ \alpha \neq 0}} \frac{q(\alpha)}{|\alpha|^2} \sum_{\substack{\varkappa \in D \\ [\varkappa/\alpha] \neq 0}} 1$$

$$= O(R^2) + O\left(\sum_{\substack{|\alpha| \leqslant R \\ \alpha \neq 0}} 1 \right) = O(R^2) \quad \text{by (4.2.2) and (4.2.3)}$$

$$\sum_{\substack{|\alpha| \leqslant R \\ \alpha \neq 0}} q(\alpha) \cdot O\left(\sum_{\substack{\varkappa \in D \\ [\varkappa/\alpha] \neq 0}} \frac{1}{|\varkappa\alpha|} \ln \left(1 + \left| \frac{\varkappa}{\alpha} \right| \right) \right)$$

$$= O\left(\sum_{\substack{|\alpha| \leqslant R \\ \alpha \neq 0}} \frac{1}{|\alpha|} \sum_{\substack{\varkappa \in D \\ [\varkappa/\alpha] \neq 0}} \frac{1}{|\varkappa|} \sum_{m \leqslant 2|\varkappa|/|\alpha|} \frac{1}{m} \right)$$

$$= O\left(\sum_{\substack{\varkappa \in D \\ \varkappa \neq 0}} \frac{1}{|\varkappa|} \sum_{m \leqslant 2|\varkappa|} \frac{1}{m} \sum_{|\alpha| \leqslant 2|\alpha|/m} \frac{1}{|\alpha|} \right)$$

$$= O\left(\sum_{\substack{\varkappa \in D \\ \varkappa \neq 0}} \frac{1}{|\varkappa|} \sum_{m \leqslant 2|\varkappa|} \frac{|\varkappa|}{m^2} \right) = O(R^2);$$

$$\sum_{\substack{\varkappa \leqslant R \\ \varkappa \neq 0}} q(\alpha) \cdot O\left(\left(\frac{Z_L}{|\alpha|} + 1 \right) \ln^2 \frac{2R}{|\alpha|} \right) = O\left(\sum_{\substack{|\alpha| \leqslant R \\ \alpha \neq 0}} \frac{R}{|\alpha|} \ln^2 \frac{2R}{|\alpha|} \right)$$

$$= O\left(\sum_{\substack{|\alpha| \leqslant R \\ \alpha \neq 0}} \sum_{m \leqslant 2R/|\alpha|} \ln^2 m \right) = O\left(\sum_{m \leqslant 2R} \ln^2 m \sum_{|\alpha| \leqslant 2R/m} 1 \right) = O(R^2).$$

As a result, we obtain

$$\sum_{\substack{|\alpha| \leqslant R \\ \alpha \neq 0}} q(\alpha) \sum_{\substack{\nu \in D/\alpha \\ \nu \neq 0}} \ln^2 |\nu| = \frac{16 \ln R}{\pi} \sum_{\varkappa \in D} 1 + O(R^2)$$

i.e.

$$\sum_{\substack{\varrho,\sigma,k,m\\ \{\varrho^k,\sigma^m\}\in D}} \ln|\varrho|\ln|\sigma| = \frac{16\ln R}{\pi}\sum_{\varkappa\in D} 1 + O(R^2).$$

And, since

$$\sum_{\substack{\varrho,\sigma,k,m\\ \{\varrho^k,\sigma^m\}\in D}} \ln|\varrho|\ln|\sigma| = 4\sum_{\varrho\in D}\ln^2|\varrho| + \sum_{\varrho,\sigma \text{ non-associated}} \ln|\varrho|\ln|\sigma| + O(R^2)$$

$$= 4\sum_{\varrho\in D}\ln^2|\varrho| + \sum_{\substack{\varrho,\sigma\\ \varrho,\sigma\in D}} \ln|\varrho|\ln|\sigma| + O(R^2)$$

Lemma 1 is proved.

We note in passing that, if we had applied the arguments of the present section to the sum

$$\sum_{\nu\in D/\alpha,\ \nu\neq 0} \ln|\nu| \quad \text{instead of} \quad \sum_{\nu\in D/\alpha,\ \nu\neq 0} \ln^2|\nu|$$

(they are correspondingly simplified), we could have obtained the asymptotic formula (4.2.5) without reference to the connection with rational primes.

4. A Recurrence Formula for the Remainder Term

It follows from (4.3.1) that

$$\sum_{\varrho\in D}\ln|\varrho| + \frac{1}{4}\sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in D}} \frac{\ln|\varrho|\ln|\sigma|}{\ln|\varrho\sigma|} = \frac{1}{\ln R}\sum_{\nu\in D}\ln^2|\varrho|$$

$$+ \frac{1}{4\ln R}\sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in D}} \ln|\varrho|\ln|\sigma| + \frac{1}{\ln R}\sum_{\varrho\in D}\ln|\varrho|\ln\frac{R}{|\varrho|}$$

$$+ \frac{1}{4\ln R}\sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in D}} \frac{\ln|\varrho|\ln|\sigma|}{\ln|\varrho\sigma|}\ln\frac{R}{|\varrho\sigma|}$$

$$= \frac{4}{\pi}\sum_{\varkappa\in D} 1 + O\left(\frac{R^2}{\ln R}\right) + O\left(\frac{1}{\ln R}\sum_{\varrho\in D}\ln|\varrho|\sum_{m\leqslant R/|\varrho|}\frac{1}{m}\right)$$

$$+ O\left(\frac{1}{\ln R}\sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in D}} \frac{\ln|\varrho|\ln|\sigma|}{\ln|\varrho\sigma|}\sum_{m\leqslant R/|\varrho\sigma|}\frac{1}{m}\right)$$

$$= \frac{4}{\pi}\sum_{\varkappa\in D} 1 + O\left(\frac{R^2}{\ln R}\right) + O\left(\frac{1}{\ln R}\sum_{m\leqslant R}\frac{1}{m}\sum_{|\varrho|\leqslant R/m}\ln|\varrho|\right)$$

$$+O\left(\frac{1}{\ln R}\sum_{m\leqslant R}\frac{1}{m}\sum_{\substack{\varrho,\sigma\\ |\varrho\sigma|\leqslant R/m}}{}'\frac{\ln|\varrho|\ln|\sigma|}{\ln|\varrho\sigma|}\right)$$

$$=\frac{4}{\pi}\sum_{\varkappa\in D}1+O\left(\frac{R^2}{\ln R}\right)+O\left(\frac{1}{\ln R}\sum_{m\leqslant R}\frac{R^2}{m^3}\right)$$

$$+O\left(\frac{1}{\ln R}\sum_{m\leqslant R}\frac{1}{m}\sum_{\substack{\varrho,\sigma\\ |\varrho\sigma|\leqslant R/m\\ |\varrho|\geqslant|\sigma|}}\ln|\sigma|\right)$$

$$=\frac{4}{\pi}\sum_{\varkappa\in D}1+O\left(\frac{R^2}{\ln R}\right)+O\left(\frac{1}{\ln R}\sum_{m\leqslant R}\frac{1}{m}\sum_{|\varrho|\leqslant\sqrt{R/m}}|\varrho|^2\right)$$

$$+O\left(\frac{1}{\ln R}\sum_{m\leqslant R}\frac{1}{m}\sum_{\sqrt{R/m}<|\varrho|\leqslant R/m}\frac{R^2}{m^2|\varrho|^2}\right)$$

$$=\frac{4}{\pi}\sum_{\varkappa\in D}1+O\left(\frac{R^2}{\ln R}\right)+O\left(\frac{1}{\ln R}\sum_{m\leqslant R}\frac{R^2}{m^3}\right)$$

$$+O\left(\frac{R^2}{\ln R}\sum_{m\leqslant R}\frac{1}{m^3}\cdot\frac{1}{\ln\frac{R}{m}}\sum_{\sqrt{R/m}<|\varrho|\leqslant R/m}\frac{\ln|\varrho|}{|\varrho|^2}\right)$$

$$=\frac{4}{\pi}\sum_{\varkappa\in D}1+O\left(\frac{R^2}{\ln R}\right)\quad\text{(we have used (4.2.5) and (4.2.8))}.$$

Thus,

$$\sum_{\varrho\in D}\ln|\varrho|+\frac{1}{4}\sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in D}}\frac{\ln|\varrho|\ln|\sigma|}{\ln|\varrho\sigma|}=\frac{4}{\pi}\sum_{\varkappa\in D}1+O\left(\frac{R^2}{\ln R}\right).\quad(4.4.1)$$

Hence (in view of the fact that (4.4.1) is applicable to the domain D/σ) it follows that

$$\sum_{|\varrho|\leqslant\sqrt{R}}\ln|\varrho|\sum_{\sigma\in D/\varrho}\ln|\sigma|+\frac{1}{4}\sum_{\substack{\varrho,\tau\\ |\varrho\tau|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\sum_{\sigma\in D/\varrho\tau}\ln|\sigma|$$

$$=\sum_{|\sigma|\leqslant R}\ln|\sigma|\left(\sum_{\substack{\varrho\in D/\sigma\\ |\tau|\leqslant\sqrt{R}}}\ln|\varrho|+\frac{1}{4}\sum_{\substack{\varrho,\tau\\ \varrho\tau\in D/\sigma\\ |\varrho\sigma|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\right)$$

$$=\sum_{\sqrt{R}<|\sigma|\leqslant R}\ln|\sigma|\left(\sum_{\varrho\in D/\sigma}\ln|\varrho|+\frac{1}{4}\sum_{\substack{\varrho,\tau\\ \varrho\tau\in D/\sigma}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\right)$$

$$+O\left(\sum_{|\sigma|\leqslant\sqrt{R}}\ln|\sigma|\left(\sum_{|\varrho|\leqslant\sqrt{R}}\ln|\varrho|+\frac{1}{4}\sum_{\substack{\varrho,\tau\\ |\varrho\tau|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\right)\right)$$

$$= \sum_{\sqrt{R}<|\sigma|\leqslant R} \left(\frac{4}{\pi} \sum_{\varkappa \in D/\sigma} 1 + O\left(\frac{R^2}{|\sigma|^2 \ln\left(\frac{R}{|\sigma|}+1\right)} \right) \right) \ln|\sigma| + O(R^2)$$

$$= \frac{4}{\pi} \sum_{\sqrt{R}<|\sigma|\leqslant R} \left(\frac{1}{|\sigma|^2} \sum_{\varkappa \in D} 1 + O\left(\frac{Z_L}{|\sigma|} + 1 \right) \right) \ln|\sigma|$$

$$+ O\left(R^2 \sum_{\sqrt{R}<|\sigma|\ R} \frac{\ln|\sigma|}{|\sigma|^2} \sum_{\ln\frac{R}{|\sigma|} < m \leqslant \ln R} \frac{1}{m^2} \right) + O(R^2)$$

$$= \frac{8 \ln R}{\pi} \sum_{\varkappa \in D} 1 + O\left(R^2 \sum_{m \leqslant \ln R} \frac{1}{m^2} \sum_{Re^{-m} < |\sigma| \leqslant R} \frac{\ln|\sigma|}{|\sigma|^2} \right) + O(R^2)$$

$$= \frac{8 \ln R}{\pi} \sum_{\varkappa \in D} 1 + O\left(R^2 \sum_{m \leqslant \ln R} \frac{1}{m} \right) + O(R^2)$$

$$= \frac{8 \ln R}{\pi} \sum_{\varkappa \in D} 1 + O(R^2 \ln\ln R)$$

(we have used (4.2.5), (4.2.12), (4.2.8), (4.2.6)).

We obtain

$$\sum_{|\varrho| \leqslant \sqrt{R}} \ln|\varrho| \sum_{\sigma \in D/\varrho} \ln|\sigma| + \frac{1}{4} \sum_{\substack{\varrho,\tau \\ |\varrho\tau| \leqslant \sqrt{R}}} \frac{\ln|\varrho| \ln|\tau|}{\ln|\varrho\tau|} \sum_{\sigma \in D/\varrho\tau} \ln|\sigma|$$

$$= \frac{8 \ln R}{\pi} \sum_{\varkappa \in D} 1 + O(R^2 \ln\ln R).$$

On multiplying this formula through by 1/4 then subtracting it from (4.3.2), we find that

$$\sum_{\varrho \in D} \ln^2|\varrho| + \frac{1}{4} \sum_{|\varrho| \leqslant \sqrt{R}} \ln|\varrho| \sum_{\sigma \in D/\varrho} \ln|\sigma| - \frac{1}{16} \sum_{\substack{\varrho,\tau \\ |\varrho\tau| \leqslant \sqrt{R}}} \frac{\ln|\varrho| \ln|\tau|}{\ln|\varrho\tau|}$$

$$\times \sum_{\sigma \in D/\varrho\tau} \ln|\sigma| = \frac{2 \ln R}{\pi} \sum_{\varkappa \in D} 1 + O(R^2 \ln\ln R). \qquad (4.4.2)$$

It is essential that the coefficients of the second and third terms on the left-hand side of (4.4.2) be in the same ratio to one another, discounting sign, as the coefficients of the first and second terms on the left-hand side of (4.4.1).

We now put

$$\sum_{\varrho \in D} \ln|\varrho| = \frac{2}{\pi} \sum_{\varkappa \in D} 1 + G(D), \qquad (4.4.3)$$

where $G(D)$ is a remainder term, the magnitude of which obviously depends on the domain D. We substitute (4.4.3) in (4.4.2):

$$\sum_{\varrho\in D}\ln^2|\varrho| = \ln R\sum_{\varrho\in D}\ln|\varrho| - \sum_{\varrho\in D}\ln|\varrho|\ln\frac{R}{|\varrho|} = \frac{2\ln R}{\pi}\sum_{\varkappa\in D}1$$

$$+\ln R\cdot G(D) + O\Big(\sum_{\varrho\in D}\ln|\varrho|\sum_{m\leqslant R/|\varrho|}\frac{1}{m}\Big) = \frac{2\ln R}{\pi}\sum_{\varkappa\in D}1 + \ln R\cdot G(D)$$

$$+O\Big(\sum_{m\leqslant R}\frac{1}{m}\sum_{|\varrho|\leqslant R/m}\ln|\varrho|\Big) = \frac{2\ln R}{\pi}\sum_{\varkappa\in D}1 + \ln R\cdot G(D) + O(R^2)$$

(we have used (4.2.5));

$$\sum_{|\varrho|\leqslant\sqrt{R}}\ln|\varrho|\sum_{\sigma\in D/\varrho}\ln|\sigma| = \frac{2}{\pi}\sum_{|\varrho|\leqslant\sqrt{R}}\ln|\varrho|\sum_{\varkappa\in D/\varrho}1$$

$$+\sum_{|\varrho|\leqslant\sqrt{R}}\ln|\varrho|\,G\Big(\frac{D}{\varrho}\Big) = \frac{2}{\pi}\sum_{|\varrho|\leqslant\sqrt{R}}\frac{\ln|\varrho|}{|\varrho|^2}\sum_{\varkappa\in D}1$$

$$+O\Big(Z_L\sum_{|\varrho|\leqslant\sqrt{R}}\frac{\ln|\varrho|}{|\varrho|^2}\Big) + O\Big(\sum_{|\varrho|\leqslant\sqrt{R}}\ln|\varrho|\Big)$$

$$+\sum_{|\varrho|\leqslant\sqrt{R}}\ln|\varrho|\,G\Big(\frac{D}{\varrho}\Big) = \frac{4\ln R}{\pi}\sum_{\varkappa\in D}1 + \sum_{|\varrho|\leqslant\sqrt{R}}\ln|\varrho|\,G\Big(\frac{D}{\varrho}\Big) + O(R^2)$$

(by (4.2.12), (4.4.1), (4.2.8), (4.2.5));

$$\sum_{\substack{\varrho,\tau\\ |\varrho\tau|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\sum_{\sigma\in D/\varrho\tau}\ln|\sigma| = \frac{2}{\pi}\sum_{\substack{\varrho,\tau\\ |\varrho\tau|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\sum_{\varkappa\in D/\varrho\tau}1$$

$$+\sum_{\substack{\varrho,\tau\\ |\varrho\tau|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\,G\Big(\frac{D}{\varrho\tau}\Big) = \frac{2}{\pi}\sum_{\substack{\varrho,\tau\\ |\varrho\tau|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{|\varrho\tau|^2\ln|\varrho\tau|}\sum_{\varkappa\in D}1$$

$$+O\Big(Z_L\sum_{\substack{\varrho,\tau\\ |\varrho\tau|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{|\varrho\tau|\ln|\varrho\tau|}\Big) + O\Big(\sum_{\substack{\varrho,\tau\\ |\varrho\tau|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\Big)$$

$$+\sum_{\substack{\varrho,\tau\\ |\varrho\tau|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\,G\Big(\frac{D}{\varrho\tau}\Big) = \frac{2}{\pi R}\sum_{\substack{\varrho,\tau\\ |\varrho\tau|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}$$

$$\times\sum_{m\leqslant R/|\varrho\tau|^2}1\sum_{\varkappa\in D}1 + O\Big(\sqrt{R}\sum_{\substack{\varrho,\tau\\ |\tau\varrho|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\sum_{m\leqslant\sqrt{R}/|\varrho\tau|}1\Big)$$

$$+O(R^2) + \sum_{\substack{\varrho,\tau\\ |\varrho\tau|\leqslant\sqrt{R}}}\frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\,G\Big(\frac{D}{\varrho\tau}\Big)$$

$$= \frac{2}{\pi R} \sum_{m \leqslant R} \sum_{\substack{\varrho, \tau \\ |\varrho\tau| \leqslant \sqrt{R/m}}} \frac{\ln|\varrho| \ln|\tau|}{\ln|\varrho\tau|} \sum_{\varkappa \in D} 1$$

$$+ O\left(\sqrt{R} \sum_{m \leqslant \sqrt{R}} \sum_{\substack{\varrho, \tau \\ |\varrho\tau| \leqslant \sqrt{R/m}}} \frac{\ln|\varrho| \ln|\tau|}{\ln|\varrho\tau|} \right) + O(R^2)$$

$$+ \sum_{\substack{\varrho, \tau \\ |\varrho\tau| \leqslant \sqrt{R}}} \frac{\ln|\varrho| \ln|\tau|}{\ln|\varrho\tau|} G\left(\frac{D}{\varrho\tau}\right)$$

$$= \frac{2}{R} \sum_{m \leqslant R} \left(\frac{16}{\pi} \sum_{|\nu| \leqslant \sqrt{R/m}} 1 - 4 \sum_{|\varrho| \leqslant \sqrt{R/m}} \ln|\varrho| + O\left(\frac{R}{m \ln\left(\frac{R}{m} + 1\right)} \right) \right) \sum_{\varkappa \in D} 1$$

$$+ O(R^2) + \sum_{\substack{\varrho, \tau \\ |\varrho\tau| \leqslant \sqrt{R}}} \frac{\ln|\varrho| \ln|\tau|}{\ln|\varrho\tau|} G\left(\frac{D}{\varrho\tau}\right)$$

$$= \frac{32 \ln R}{\pi} \sum_{\varkappa \in D} 1 - \frac{8}{\pi R} \sum_{|\varrho| \leqslant \sqrt{R}} \ln|\varrho| \sum_{m \leqslant R/|\varrho^2|} 1 \sum_{\varkappa \in D} 1$$

$$+ O\left(R^2 \sum_{\sqrt{R} < m \leqslant R} \frac{1}{m} \sum_{\ln \frac{R}{m} < n \leqslant \ln R} \frac{1}{n^2} \right) + O(R^2)$$

$$+ \sum_{\substack{\varrho, \tau \\ |\varrho\tau| \leqslant \sqrt{R}}} \frac{\ln|\varrho| \ln|\tau|}{\ln|\varrho\tau|} G\left(\frac{D}{\varrho\tau}\right) = \frac{32 \ln R}{\pi} \sum_{\varkappa \in D} 1 - \frac{8}{\pi} \sum_{|\varrho| \leqslant \sqrt{R}} \frac{\ln|\varrho|}{|\varrho|^2}$$

$$\times \sum_{\varkappa \in D} 1 + O\left(R \sum_{|\varrho| \leqslant \sqrt{R}} \ln|\varrho| \right) + O\left(R^2 \sum_{n \leqslant \ln R} \frac{1}{n^2} \sum_{Re^{-n} < m \leqslant R} \frac{1}{m} \right)$$

$$+ O(R^2) + \sum_{\substack{\varrho, \tau \\ |\varrho\tau| \leqslant \sqrt{R}}} \frac{\ln|\varrho| \ln|\tau|}{\ln|\varrho\tau|} G\left(\frac{D}{\varrho\tau}\right) = \frac{16 \ln R}{\pi} \sum_{\varkappa \in D} 1$$

$$+ O\left(R^2 \sum_{n \leqslant \ln R} \frac{1}{n} \right) + O(R^2) + \sum_{\substack{\varrho, \tau \\ |\varrho\tau| \leqslant \sqrt{R}}} \frac{\ln|\varrho| \ln|\tau|}{\ln|\varrho\tau|} G\left(\frac{D}{\varrho\tau}\right)$$

$$= \frac{16 \ln R}{\pi} \sum_{\varkappa \in D} 1 + \sum_{\substack{\varrho, \tau \\ |\varrho\tau| \leqslant \sqrt{R}}} \frac{\ln|\varrho| \ln|\tau|}{\ln|\varrho\tau|} G\left(\frac{D}{\varrho\tau}\right) + O(R^2 \ln \ln R)$$

(by (4.2.12), (4.4.1), (4.2.8), (4.2.5)),

After substituting the expressions obtained for

$$\sum_{|\varrho|<R} \ln^2|\varrho|,\quad \sum_{|\varrho|<\sqrt{R}} \ln|\varrho| \sum_{\sigma\in D/\varrho} \ln|\sigma|,\quad \sum_{\substack{\varrho,\tau\\ |\varrho\tau|<\sqrt{R}}} \frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|} \sum_{\sigma\in D/\varrho\tau} \ln|\sigma|$$

in (4.4.2) and cancelling, we get

$$\ln R\cdot G(D) + \frac{1}{4}\sum_{|\varrho|<\sqrt{R}} \ln|\varrho|\, G\left(\frac{D}{\varrho}\right) - \frac{1}{16}\sum_{\substack{\varrho,\tau\\ |\varrho\tau|<\sqrt{R}}} \frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|} \times G\left(\frac{D}{\varrho\tau}\right) = O(R^2\ln\ln R),$$

i.e.

$$G(D) = -\frac{1}{4\ln R}\sum_{|\varrho|<\sqrt{R}} \ln|\varrho|\, G\left(\frac{D}{\varrho}\right) + \frac{1}{16\ln R}\sum_{\substack{\varrho,\tau\\ |\varrho\tau|<\sqrt{R}}} \frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|} G\left(\frac{D}{\varrho\tau}\right) + O\left(\frac{R^2\ln\ln R}{\ln R}\right).$$

An immediate consequence of this is the lemma.

LEMMA 2. *The asymptotic recurrence formula*

$$|G(D)| \leqslant \frac{1}{4\ln R}\sum_{|\varrho|<\sqrt{R}} \ln|\varrho|\cdot\left|G\left(\frac{D}{\varrho}\right)\right| + \frac{1}{16\ln R}\sum_{\substack{\varrho,\tau\\ |\varrho\tau|<\sqrt{R}}} \frac{\ln|\varrho|\ln|\tau|}{\ln|\varrho\tau|}\left|G\left(\frac{D}{\varrho\tau}\right)\right| + O\left(\frac{R^2\ln\ln R}{\ln R}\right) \qquad (4.4.4)$$

holds for $G(D)$.

This recurrence formula, in conjunction with the arguments of the next section, provides the basis for the proof of the law of distribution of Gaussian primes.

5. "Islets" with Small Values of $\left|G\left(\frac{D}{\nu}\right)\right|$

Our problem is to verify that there exists in the circle $|\xi| \leqslant R$ a sufficiently large number of sufficiently large circles such that, for ν lying inside them, $|G(D/\nu)|$ is sufficiently small.

We approach this problem by considering the sum $\sum_{z<|\nu|\leqslant y} G(D/\nu)$, where $1 \leqslant z < y \leqslant R$; we start with the sum $\sum_{1\leqslant|\nu|\leqslant y} G(D/\nu)$:

$$\sum_{1\leqslant|\nu|\leqslant y} G\left(\frac{D}{\nu}\right) = \sum_{1\leqslant|\nu|\leqslant y}\left(\sum_{\varrho\in D/\nu}\ln|\varrho| - \frac{2}{\pi}\sum_{\varkappa\in D/\nu} 1\right)$$

$$= \sum_{|\varrho|\leqslant R}\ln|\varrho| \sum_{\substack{1\leqslant|\nu|\leqslant y\\ \nu\in D/\varrho}} 1 - \frac{2}{\pi}\sum_{1\leqslant|\nu|<y}\left(\frac{1}{|\nu|^2}\sum_{\varkappa\in D} 1 + O\left(\frac{Z_L}{|\nu|}\right) + O(1)\right)$$

$$= \sum_{|\varrho|\leqslant R/y}\ln|\varrho|\,.\,O(y^2) + \sum_{R/y<|\varrho|\leqslant R}\ln|\varrho| \sum_{\substack{\nu\in D/\varrho\\ \nu\neq 0}} 1$$

$$-\frac{2}{\pi}\sum_{1\leqslant|\nu|\leqslant y}\frac{1}{|\nu|^2}\sum_{\varkappa\in D} 1 + O(Z_L y) + O(y^2) = \sum_{R/y<|\varrho|\leqslant R}\ln|\varrho|$$

$$\times\left(\frac{1}{|\varrho|^2}\sum_{\varkappa\in D} 1 + O\left(\frac{Z_L}{|\varrho|}\right) + O(1)\right) - 4\ln y\sum_{\varkappa\in D} 1 + O(R^2)$$

$$= 4\ln y\sum_{\varkappa\in D} 1 + O(R^2) - 4\ln y\sum_{\varkappa\in D} 1 + O(R^2) = O(R^2)$$

(by (4.4.3), (4.2.12), (4.2.10), (4.2.5), (4.5.8) and (4.2.6))

But now, we also have

$$\sum_{z<|\nu|\leqslant y} G\left(\frac{D}{\nu}\right) = O(R^2). \tag{4.5.1}$$

We shall also require an asymptotic formula for $G(D/\nu) - G(D/\mu)$. Let $|\nu| \leqslant R$, $|\mu| \leqslant R$, $\min(|\nu|, |\mu|) = h$, and let the domain $\Delta = D/\nu + D/\mu$. The condition imposed on D, which also holds for D/ν and D/μ, may easily be seen to apply likewise to Δ. Formula (4.4.1) is therefore applicable to all three of these domains, as also to $\Delta - D/\nu$ and $\Delta - D/\mu$. It follows from

$$\sum_{\varrho\in\Delta-D/\nu}\ln|\varrho| = \frac{4}{\pi}\sum_{\varkappa\in\Delta-D/\nu} 1 - \frac{1}{4}\sum_{\substack{\varrho,\sigma\\ \varrho\sigma\in\Delta-D/\nu}}\frac{\ln|\varrho|\ln|\sigma|}{\ln|\varrho\sigma|}$$

$$+ O\left(\frac{R^2}{h^2\ln\left(\frac{R}{h}+1\right)}\right)$$

that

$$0 \leqslant \sum_{\varrho\in D}\ln|\varrho| - \sum_{\varrho\in D/\nu}\ln|\varrho| \leqslant \frac{4}{\pi}\sum_{\substack{\varkappa\in\Delta\\ \varkappa\notin D/\nu}} 1 + O\left(\frac{R^2}{h^2\ln\left(\frac{R}{h}+1\right)}\right)$$

(the left-hand inequality is obvious).

Similarly,

$$0 \leqslant \sum_{\varrho \in \Delta} \ln|\varrho| - \sum_{\varrho \in D/\mu} \ln|\varrho| \leqslant \frac{4}{\pi} \sum_{\substack{x \in \Delta \\ x \notin D/\mu}} 1 + O\left(\frac{R^2}{h^2 \ln\left(\frac{R}{h}+1\right)}\right).$$

Consequently,

$$\sum_{\varrho \in D/\nu} \ln|\varrho| - \sum_{\varrho \in D/\mu} \ln|\varrho| \begin{cases} \leqslant \frac{4}{\pi} \sum\limits_{\substack{x \in \Delta \\ x \notin D/\mu}} 1 + O\left(\frac{R^2}{h^2 \ln\left(\frac{R}{h}+1\right)}\right), \\ \geqslant -\frac{4}{\pi} \sum\limits_{\substack{x \in \Delta \\ x \notin D/\nu}} 1 + O\left(\frac{R^2}{h^2 \ln\left(\frac{R}{h}+1\right)}\right). \end{cases}$$

But in this case, by (4.4.3),

$$\left|G\left(\frac{D}{\nu}\right) - G\left(\frac{D}{\mu}\right)\right| \leqslant \frac{2}{\pi} \sum_{\substack{x \in \Delta \\ x \notin D/\nu}} 1 + \frac{2}{\pi} \sum_{\substack{x \in \Delta \\ x \notin D/\mu}} 1 + O\left(\frac{R^2}{h^2 \ln\left(\frac{R}{h}+1\right)}\right)$$

$$= \frac{2}{\pi} \sum_{\substack{x \in D/\mu \\ x \notin D/\nu}} 1 + \frac{2}{\pi} \sum_{\substack{x \in D/\nu \\ x \notin D/\mu}} 1 + O\left(\frac{R^2}{h^2 \ln\left(\frac{R}{h}+1\right)}\right).$$

But it is easily seen that, if $\xi \in D/\mu$ and $\xi \notin D/\nu$, then there exists $\xi_1 \in D/\nu$ such that $|\xi - \xi_1| \leqslant R|\mu - \nu|/|\mu\nu|$ (it is sufficient to take $\xi_1 = \xi\mu/\nu$), and also a $\xi_2 \notin D/\mu$ such that $|\xi - \xi_2| \leqslant R|\mu - \nu|/|\mu|^2$ (it is sufficient to take $\xi_2 = \xi\nu/\mu$). The situation is similar for $\xi \in D/\nu$, $\xi \notin D/\mu$. On applying formulae (4.2.11) and (4.2.13) to L/μ if $|\mu| \geqslant |\nu|$, and to L/ν if $|\nu| \geqslant |\mu|$, we obtain

$$\sum_{\substack{x \in D/\mu \\ x \notin D/\nu}} 1 + \sum_{\substack{x \in D/\nu \\ x \notin D/\mu}} 1 \ll R^2 \frac{|\mu - \nu|\,h}{|\mu\nu|^2} + R^2 \frac{|\mu - \nu|^2}{|\mu\nu|^2} + \frac{Rh}{|\mu\nu|},$$

whence

$$\left|G\left(\frac{D}{\nu}\right) - G\left(\frac{D}{\mu}\right)\right| \ll R^2 \frac{|\mu - \nu|\,h}{|\mu\nu|^2} + R^2 \frac{|\mu - \nu|^2}{|\mu\nu|^2} + \frac{R^2}{h^2 \ln\left(\frac{R}{h}+1\right)}. \tag{4.5.2}$$

If $|\mu| \asymp |\nu|$ and $|\mu - \nu| = O(1)$, then

$$\left|G\left(\frac{D}{\nu}\right) - G\left(\frac{D}{\mu}\right)\right| \ll \frac{R^2}{h^2} + \frac{R^2}{h^2 \ln\left(\frac{R}{h}+1\right)}. \tag{4.5.3}$$

We now take an arbitrary $s > 0$ and show that, provided certain conditions are imposed on y and z, there exists a v_0 in the ring $z < |v| \leqslant y$ such that

$$\left|G\left(\frac{D}{v_0}\right)\right| \leqslant \frac{sR^2}{|v_0|^2}.$$

We consider two cases here:

(1) When $z < |v| \leqslant y$ all the $G(D/v)$ have the same sign. Now,

$$\left|\sum_{z<|v|\leqslant y} G\left(\frac{D}{v}\right)\right| = \sum_{z<|v|\leqslant y}\left|G\left(\frac{D}{v}\right)\right|,$$

and (4.5.1) yields

$$\sum_{z<|v|\leqslant y}\left|G\left(\frac{D}{v}\right)\right| = O(R^2).$$

And since, with $\ln (y/z) \geqslant c_3/z$ (where $c_3 > 0$),

$$\sum_{z<|v|\leqslant y}\left|G\left(\frac{D}{v}\right)\right| \geqslant \min_{z<|v|\leqslant y} \frac{|v|^2\left|G\left(\frac{D}{v}\right)\right|}{R^2} \times \sum_{z<|v|\leqslant y}\frac{R^2}{|v|^2} \gg \min_{z<|v|\leqslant y} \frac{|v|^2\left|G\left(\frac{D}{v}\right)\right|}{R^2} R^2 \ln\frac{y}{z},$$

then (when $\ln y/z \gg c_3/z$)

$$\min_{z<|v|\leqslant y} \frac{|v|^2\left|G\left(\frac{D}{v}\right)\right|}{R^2} < \frac{c_4}{\ln\frac{y}{z}},$$

where $c_4 > 0$. In other words, given the supplementary condition $\ln (y/z) \geqslant c_4/s$, there exists v_0 such that $z < |v_0| \leqslant y$ and $|G(D/v_0)| \leqslant sR^2/|v_0|^2$.

(2) When $z < |v| \leqslant y$, there are $G(D/v)$ of different signs: $G(D/v') < 0$, $G(D/v'') > 0$. When $y - z \geqslant 1$, we can choose a sequence of numbers $v_1 = v', v_2, v_3, \ldots, v_{n-1}, v_n = v''$, satisfying the condition $z < |v_t| \leqslant y$ and such that $|v_t - v_{t+1}| = O(1)$, $|v_t| \rightleftharpoons |v_{t+1}|$ $(t = 1, 2, \ldots, n-1)$ (obviously, n is not necessarily negative). There exists t $(1 \leqslant t \leqslant n-1)$ such that

$$G\left(\frac{D}{v_t}\right) \leqslant 0, \quad G\left(\frac{D}{v_{t+1}}\right) > 0.$$

Having chosen any one of ν_t, ν_{t+1} as ν_0, we shall have, in accordance with (4.5.3):

$$\left|G\left(\frac{D}{\nu_0}\right)\right| \leqslant \left|G\left(\frac{D}{\nu_t}\right) - G\left(\frac{D}{\nu_{t+1}}\right)\right| \leqslant \frac{R^2}{|\nu_0|^3} + \frac{R^2}{|\nu_0|^2 \ln\left(\frac{R}{|\nu_0|} + 1\right)}.$$

It is clear from this that constants $c_5 > 0$, $c_6 > 0$ can be chosen such that, provided $z > c_5/s$, $y \leqslant Re^{-c_6/s}$,

$$\left|G\left(\frac{D}{\nu_0}\right)\right| \leqslant \frac{sR^2}{|\nu_0|^2}.$$

Thus the inequality is satisfied in both cases.

Now let $z = u$, $y = ue^{c_4/s}$. Let $u > \frac{c_5}{s}$, $u > \frac{c_3 + 1}{c_4} s$,

$$u \leqslant Re^{-\frac{c_6 + c_4}{s}}.$$

In this case all the conditions

$$\ln\frac{y}{z} \geqslant \frac{c_3}{z}, \quad \ln\frac{y}{z} \geqslant \frac{c_4}{s}, \quad y - z \geqslant 1, \quad z > \frac{c_5}{s}, \quad y \leqslant Re^{-c_6/s}$$

will be satisfied and, whether the ring $u < |\nu| \, ue^{c_4/s}$ contains $G(D/\nu)$ of different signs or not, there is a ν_0 in this ring such that

$$\left|G\left(\frac{D}{\nu_0}\right)\right| \leqslant \frac{sR^2}{|\nu_0|^2}. \tag{4.5.4}$$

We shall bound s from above: $s \leqslant c_7$ (the choice of c_7 is at our disposal). Let $|\nu - \nu_0| \leqslant s|\nu_0|/c_8$, where $c_8 \geqslant 2c_7$. Now, $|\nu - \nu_0| \leqslant \frac{1}{2}|\nu_0|$, $|\nu| \rightleftharpoons |\nu_0|$ and formula (4.5.2) gives, with $u = \nu_0$:

$$\left|G\left(\frac{D}{\nu}\right) - G\left(\frac{D}{\nu_0}\right)\right| \ll R^2 \frac{|\nu - \nu_0|}{|\nu_0|^3} + R^2 \frac{|\nu - \nu_0|^2}{|\nu_0|^4} + \frac{R^2}{|\nu_0|^2 \ln\left(\frac{R}{|\nu_0|} + 1\right)}.$$

Hence it follows that, if the constant c_8 in $|\nu - \nu_0| \leqslant s|\nu_0|/c_8$ is sufficiently large, then

$$\left|G\left(\frac{D}{\nu}\right) - G\left(\frac{D}{\nu_0}\right)\right| \leqslant \frac{sR^2}{|\nu_0|^2}. \tag{4.5.5}$$

It follows from (4.5.4), (4.5.5) and $|\nu_0| \geqslant 2|\nu|/3$ (since $|\nu - \nu_0| \leqslant \frac{1}{2}|\nu_0|$) that

LEMMA 3.

If $$u > \frac{c_5}{s}, \quad u > \frac{c_3+1}{c_4}s, \quad u \leqslant Re^{-\frac{c_6+c_4}{s}},$$

then there exists a v_0 $(u < |v_0| \leqslant uec_4/s)$ *such that, when* $|v - v_0| \leqslant s|v_0|/c$,

$$\left|G\left(\frac{D}{v}\right)\right| \leqslant \frac{9sR^2}{2|v|^2}. \tag{4.5.6}$$

6. Proof of the Theorem

LEMMA 4.

$$G(D) = O\left(\frac{R^2}{\sqrt[3]{\ln\ln R}}\right). \tag{4.6.1}$$

We shall prove this by induction. Let $R \geqslant R_0 \geqslant 300$; suppose that, for all v $(1 < |v| \leqslant \sqrt{R})$,

$$\left|G\left(\frac{D}{v}\right)\right| \leqslant \frac{BR^2}{|v|^2\sqrt[3]{\ln\ln\frac{R}{|v|}}} \leqslant \frac{BR^2}{|v|^2\sqrt[3]{\ln\ln\sqrt{R}}} \tag{4.6.2}$$

(where $B \geqslant 1$).

It is obvious in general that $G(D/v) \ll R^2/|v|^2$, i.e. $|G(D/v)| \ll c_9R^2/|v|^2$. We equate the arbitrary constant c_7 to c_9. On the assumption that

$$\frac{B}{9\sqrt[3]{\ln\ln\sqrt{R_0}}} \leqslant c_9,$$

we put in Lemma 3:

$$s = \frac{B}{9\sqrt[3]{\ln\ln\sqrt{R}}}.$$

The numbers

$$u_n\left(0 \leqslant n \leqslant \frac{s\ln R}{2c_4} - \frac{s\ln u_0}{c_4}\right)$$

are defined thus:

$$u_0 = 9c_5\sqrt[3]{\ln\ln\sqrt{R}} + \frac{(c_3+1)B}{9c_4\sqrt[3]{\ln\ln\sqrt{R}}} + 1;$$

$$u_n = u_0e^{nc_4/s} \text{ (so that } u_ne^{c_4/s} = v_{n+1}).$$

When

$$0 \leqslant n \leqslant t = \frac{s \ln R}{2c_4} - \frac{s \ln u_0}{c_4} - 1$$

and

$$18c_6 \sqrt[3]{\ln \ln \sqrt{R_0}} \leqslant \ln R_0$$

the conditions

$$u_n > \frac{c_5}{s}, \quad u_n > \frac{c_3 + 1}{c_4} s, \quad u_n \leqslant \mathrm{Re}^{-\frac{c_0 + c_4}{s}}$$

will be satisfied, and by Lemma 3, there is a $\nu_{0,n}$ ($u_n < |\nu_{0,n}| \leqslant u_{n+1}$) such that, for all ν satisfying $|\nu - \nu_{0,n}| \leqslant s|\nu_{0,n}|/c_8$, we have

$$\left| G\left(\frac{D}{\nu}\right) \right| \leqslant \frac{9sR^2}{2|\nu|^2} = \frac{BR^2}{2|\nu|^2 \sqrt[3]{\ln \ln \sqrt{R}}}. \tag{4.6.3}$$

Using (4.4.4), (4.6.2), (4.6.3), (4.4.1), we have:

$$\begin{aligned}
G(D)| &\leqslant \frac{1}{4 \ln R} \sum_{|\varrho| < \sqrt{R}} \ln |\varrho| \left| G\left(\frac{D}{\varrho}\right) \right| \\
&+ \frac{1}{16 \ln R} \sum_{\substack{\varrho, \tau \\ |\varrho\tau| < \sqrt{R}}} \frac{\ln |\varrho| \ln |\tau}{\ln |\varrho\tau|} \left| G\left(\frac{D}{\varrho\tau}\right) \right| + O\left(\frac{R^2 \ln \ln R}{\ln R}\right) \\
&\leqslant \frac{BR}{4 \ln R \sqrt[3]{\ln \ln \sqrt{R}}} \left(\sum_{|\varrho| < \sqrt{R}} \frac{R}{|\varrho|^2} \ln |\varrho| + \frac{1}{4} \sum_{\substack{\varrho, \tau \\ |\varrho\tau| < \sqrt{R}}} \frac{R \ln |\varrho| \ln \tau}{|\varrho\tau|^2 \ln |\varrho\tau|} \right) \\
&- \frac{BR^2}{8 \ln R \sqrt[3]{\ln \ln \sqrt{R}}} \sum_{0 < n < t} \left(\sum_{\substack{u_n < |\varrho| \leqslant u_{n+1} \\ |\varrho - \nu_{0,n}| < \frac{s|\nu_{0,n}|}{c_8}}} \frac{\ln |\varrho|}{|\varrho|^2} \right. \\
&\left. + \frac{1}{4} \sum_{\substack{\varrho, \tau \\ u_n < |\varrho\tau| \leqslant u_{n+1} \\ |\varrho\tau - \nu_{0,n}| < \frac{s|\nu_{0,n}|}{c_8}}} \frac{\ln |\varrho| \ln |\tau|}{|\varrho\tau|^2 \ln |\varrho\tau|} \right) + O\left(\frac{R^2 \ln \ln R}{\ln R}\right) \\
&\leqslant \frac{BR}{4 \ln R \sqrt[3]{\ln \ln \sqrt{R}}} \left(\sum_{|\varrho| < \sqrt{R}} \ln |\varrho| \sum_{m < \frac{1}{|\varrho|^2}} 1 \right. \\
&\left. + \frac{1}{4} \sum_{\substack{\varrho, \tau \\ |\varrho\tau| < \sqrt{R}}} \frac{\ln |\varrho| \ln |\tau|}{\ln |\varrho\tau|} \sum_{m < \frac{1}{|\varrho\tau|^2}} 1 \right) -
\end{aligned}$$

$$
\begin{aligned}
&- \frac{BR^2}{8 \ln R \sqrt[3]{\ln \ln \sqrt{R}}} \sum_{0 \leqslant n < t} \frac{4}{9|\nu_{0,n}|^2} \left(\sum_{\substack{u_n < |\varrho| \leqslant u_{n+1} \\ |\varrho - \nu_{0,n}| \leqslant \frac{s|\nu_{0,n}|}{c_8}}} \ln |\varrho| \right. \\
&\left. + \frac{1}{4} \sum_{\substack{\varrho, \tau \\ u_n < |\varrho\tau| \leqslant |u_{n+1} \\ |\varrho\tau - \nu_{0,n}| \leqslant \frac{s|\nu_{0,n}|}{c_8}}} \frac{\ln |\varrho| \ln |\tau|}{\ln |\varrho\tau|} \right) + O\left(\frac{R^2 \ln \ln R}{\ln R} \right) \\
&= \frac{BR}{4 \ln R \sqrt[3]{\ln \ln \sqrt{R}}} \\
&\times \sum_{m \leqslant R} \left(\sum_{|\varrho| \leqslant \sqrt{R/m}} \ln |\varrho| + \frac{1}{4} \sum_{|\varrho\tau| \leqslant \sqrt{R/m}} \frac{\ln |\varrho| \ln |\tau|}{\ln |\varrho\tau|} \right) \\
&- \frac{BR^2}{18 \ln R \sqrt[3]{\ln \ln \sqrt{R}}} \sum_{0 \leqslant n < t} \frac{1}{|\nu_{0,n}|^2} \\
&\times \left(\frac{4}{\pi} \sum_{\substack{u_n < |\varkappa| \leqslant u_{n+1} \\ |\varkappa - \nu_{0,n}| \leqslant \frac{s|\nu_{0,n}|}{c_8}}} 1 + O\left(\frac{|\nu_{0,n}|^2}{\ln |\nu_{0,n}|} \right) \right) + O\left(\frac{R^2 \ln \ln R}{\ln R} \right).
\end{aligned}
$$

Since $s|\nu_{0,n}|/c_8 \leqslant \frac{1}{2}|\nu_{0,n}|$, it is obvious that

$$
\sum_{\substack{u_n < |\varkappa| \leqslant u_{n+1} \\ |\varkappa - \nu_{0,n}| \leqslant \frac{s|\nu_{0,n}|}{c_8}}} 1 \geqslant \frac{\pi c_{10} s^2 |\nu_{0,n}|^2}{c_8^2} + O(s|\nu_{0,n}|) \qquad (0 < c_{10} < 1).
$$

We continue, to obtain

$$
\begin{aligned}
|G(D)| &\leqslant \frac{BR}{4 \ln R \sqrt[3]{\ln \ln \sqrt{R}}} \\
&\times \sum_{m \leqslant R} \left(\frac{4}{\pi} \sum_{|\varkappa| \leqslant \sqrt{R/m}} 1 + O\left(\frac{R}{m \ln\left(\frac{R}{m} + 1 \right)} \right) \right) \\
&- \frac{BR^2}{8 \ln R \sqrt[3]{\ln \ln \sqrt{R}}} \sum_{0 \leqslant n < t} \left(\frac{4 c_{10} s^2}{c_8^2} + O\left(\frac{s}{|\nu_{0,n}|} \right) \right. \\
&\left. + O\left(\frac{1}{\ln |\nu_{0,n}|} \right) \right) + O\left(\frac{R^2 \ln \ln R}{\ln R} \right) = \frac{BR^2}{\sqrt[3]{\ln \ln \sqrt{R}}}
\end{aligned}
$$

$$-\frac{2c_{10}BR^2s^2t}{9c_8^2\ln R\sqrt[3]{\ln\ln\sqrt{R}}}+\frac{BR^2}{18\ln R\sqrt[3]{\ln\ln\sqrt{R}}}$$
$$\times O\Big(\sum_{0\leqslant n\leqslant t}\frac{1}{n+1}\Big)+O\Big(\frac{R^2\ln\ln R}{\ln R}\Big)=\frac{BR^2}{\sqrt[3]{\ln\ln\sqrt{R}}}$$
$$-\frac{c_{10}B^4R^2}{6561c_4c_8^2(\ln\ln\sqrt{R})^{4/3}}+O\Big(\frac{R^2\ln u_0}{\ln R}\Big)+O\Big(\frac{R^2\ln\ln R}{\ln R}\Big).$$

But it is clear from the definition of u_0 that

$$u_0=9c_5\sqrt[3]{\ln\ln\sqrt{R}}+O(1).$$

Our working therefore yields

$$|G(D)|\leqslant\frac{BR^2}{\sqrt[3]{\ln\ln\sqrt{R}}}-\frac{c_{10}B^4R^2}{6561c_4c_8^2(\ln\ln\sqrt{R})^{4/3}}+O\Big(\frac{R^2\ln\ln R}{\ln R}\Big).$$

Consequently, when $R_0\geqslant c_{11}$,

$$|G(D)|\leqslant\frac{BR^2}{\sqrt[3]{\ln\ln\sqrt{R}}}-\frac{c_{12}B^4R^2}{(\ln\ln\sqrt{R})^{4/3}}\leqslant\frac{BR^2}{\sqrt[3]{\ln\ln R}}$$
$$+\frac{c_{13}BR^2}{(\ln\ln\sqrt{R})^{4/3}}-\frac{c_{12}B^4R^2}{(\ln\ln\sqrt{R})^{4/3}}\qquad(c_{12}>0,\ c_{13}>0).$$

Finally, it follows from this that, when $B^3>c_{13}/c_{12}$,

$$|G(D)|\leqslant\frac{BR^2}{\sqrt[3]{\ln\ln R}}.$$

We now choose constant values for B, R_0 such that the following conditions are satisfied:

$$B\geqslant1,\quad B^3\geqslant\frac{c_{13}}{c_{12}},\quad R_0\geqslant300,\quad R_0\geqslant c_{11},\quad 18c_6\sqrt[3]{\ln\ln\sqrt{R_0}}\leqslant\ln R_0,$$
$$c_9\leqslant\frac{B}{\sqrt[3]{\ln\ln\sqrt{R_0}}},\quad\frac{B}{9\sqrt[3]{\ln\ln\sqrt{R_0}}}\leqslant c_9.$$

This can obviously be done. Now, when $R\leqslant R_0$, we have

$$|G(D)|\leqslant c_9R^2\leqslant\frac{BR^2}{\sqrt[3]{\ln\ln R_0}}\leqslant\frac{BR^2}{\sqrt[3]{\ln\ln R}}.$$

According to our calculations, the inequality $|G(D)| \leqslant BR^2/\sqrt[3]{\ln\ln R}$ will now also be satisfied with $R \leqslant R_0\sqrt{2}$ (since $R \leqslant R_0\sqrt{2}$ and $|\nu| > 1$ imply $R/|\nu| \leqslant R_0$, and the inequality is fulfilled for $|G(D/\nu)|$). But in this case it also holds for $R \leqslant 2R_0$, i.e. also for $R \leqslant 2R_0\sqrt{2}$ and so on, and hence for any R. Since B is constant, lemma 4 is proved.

Our theorem follows immediately from Lemma 4. For,

$$\sum_{\varrho\in D} 1 = \frac{1}{\ln R}\sum_{\varrho\in D}\ln|\varrho| + \frac{1}{\ln R}\sum_{\varrho\in D}\ln\frac{R}{|\varrho|} = \frac{2}{\pi\ln R}\sum_{\varkappa\in D} 1$$
$$+ \frac{G(D)}{\ln R} + O\left(\frac{1}{\ln R}\sum_{\varrho\in D}\sum_{m\leqslant R/|\varrho|}\frac{1}{m}\right) = \frac{2}{\pi\ln R}\sum_{\varkappa\in D} 1$$
$$+ O\left(\frac{R^2}{\ln R\sqrt[3]{\ln\ln R}}\right) + O\left(\frac{1}{\ln R}\sum_{m\leqslant R}\frac{1}{m}\sum_{|\varrho|\leqslant R/m} 1\right) = \frac{2}{\pi\ln R}\sum_{\varkappa\in D} 1$$
$$+ O\left(\frac{R^3}{\ln R\sqrt[3]{\ln\ln R}}\right) + O\left(\frac{R^2}{\ln R}\sum_{m\leqslant R}\frac{1}{m^3\ln\left(\frac{R}{m}+1\right)}\right)$$
$$= \frac{2}{\pi\ln R}\sum_{\varkappa\in D} 1 + O\left(\frac{R^2}{\ln R\sqrt[3]{\ln\ln R}}\right),$$

and the theorem is proved.

7. A Theorem on Almost-prime Gaussian Numbers

We shall describe a number as almost-prime if it has only a limited number of divisors. The properties of such numbers thus approximate to those of primes. It is natural to regard numbers of the form pp', where p, p' are primes, as the closest to primes in their properties.

It can be deduced, by using the methods of Chapter 5, that irreducible numerical polynomials take almost-prime values infinitely many times; whether or not they take prime values infinitely many times is as yet unknown.

In particular, it is unknown whether or not there exist infinitely many primes of the form $p = x^2 + 1$.

Obviously, this problem is equivalent to the problem of whether or not infinitely many Gaussian primes of the form $x + i$ exist.

An approach to this problem is provided by the question of the existence of Gaussian primes of the form $x + iy$, where $|y| \leqslant \psi(x)$, and $\psi(x)$ is a slowly increasing function as $x \to \infty$. For instance, it can be proved by non-elementary analytical methods (accepting Riemann's hypothesis for Hecke series of the field $k(\sqrt{-1})$) that there exist infinitely many primes of the above-mentioned form with $\psi(x) = O[(\ln x)^{3/2}]$. If, instead of Gaussian primes, we consider almost prime numbers of the form $\varrho\varrho'$, where ϱ, ϱ' are Gaussian primes, a theorem of a similar kind can be proved about them by purely elementary means (see I. P. Kubilyus and Yu. V. Linnik[10]).

THEOREM 4.7.1 (Kubilyus and Linnik). *There exist infinitely many pairs of primes pp' such that*

$$pp' = X^2 + y^2, \tag{4.7.1}$$

where

$$y = O(\ln X). \tag{4.7.2}$$

For the proof, we consider the ring $\sqrt{R} < |z| < R$ on the complex plane of Gaussian numbers. The number of Gaussian primes ϱ under the condition $\sqrt{R} < |\varrho| < R$ will be not less than $c(R^2/(\ln R)$ in this ring (see § 1 of this chapter, where it becomes a matter of computing the primes $p \equiv 1 \pmod 4$; $p < R^2$). For each such number ϱ_j we write $|\varrho_j| = \varrho\, r_i$ and $\arc \varrho_j = \varphi_j$. Obviously, there exist at least two number ϱ_{j1} and ϱ_{j2} in our ring such that

$$|\arc \varrho_{j_2} - \arc \varrho_{j_1}| = \left|\arc \frac{\varrho_{j_2}}{\varrho_{j_1}}\right| = B\frac{\ln R}{R^2}.$$

We form the number $\varrho_{j2}\bar{\varrho}_{j1}$. We have

$$\arc \varrho_{j_2}\bar{\varrho}_{j_1} = B\frac{\ln R}{R^2}; \quad |\varrho_{j_2}\bar{\varrho}_{j_1}| = BR^2.$$

Hence

$$\varrho_{j_2}\bar{\varrho}_{j_1} = X + iy, \quad y = B\ln R = B\ln X,$$

which proves our theorem.

CHAPTER 5

THE SIEVE OF ERATOSTHENES

THE principle of the "sieve" of Eratosthenes has been familiar since classical times. It was used for computing the number of primes in a given interval. In 1919 Viggo Brun[33] and in 1946 Atle Selberg[45] proposed new forms of this method for estimating the number of primes in two sequences at once. Chapters 5 and 6 will be devoted to a discussion of these elementary principles.

1. The "Double Rectangular Sieve"

Before discussing Brun's ideas, we shall describe the form of a double sieve by taking a concrete example. The word "double" will imply in future that the "sieve" method is applied to two sequences at once. As our example, let us prove the following theorem.

THEOREM 5.1.1. *The series consisting of the reciprocals of twin primes is convergent.*

Let $\pi_2(x)$ denote the number of pairs of twin primes each of which does not exceed x. The method of proof consists in obtaining a non-trivial upper bound for $\pi_2(x)$. After this, the theorem is easily proved by a partial summation.

We shall start by proving:

THEOREM 5.1.2. *The inequality*

$$\pi_2(x) < c \frac{x}{\ln^2 x} (\ln \ln x)^2, \tag{5.1.1}$$

where c, x_0 are positive absolute constants, holds for $\pi_2(x)$ when $x > x_0$.

We shall "sift" all the numbers in the sequence $a_n = n(n+2)$, which are divisible by at least one prime $\leqslant z$, where $z < \sqrt{x}$, $n \leqslant x$. We denote by R the product of all the primes $\leqslant z$. If we

stop at the even step $2k$ in the "sieve" method, we get the inequality

$$\pi_2(x) \leqslant x - \sum_{p \mid R} \sum_{p \mid a_n} 1 + \sum_{p_1 p_2 \mid R} \sum_{p_1 p_2 \mid a_n} 1 - \\ \ldots + \sum_{p_1 \ldots p_{2k} \mid R} \sum_{p_1 \ldots p_{2k} \mid a_n} 1. \tag{5.1.2}$$

The numbers z and k are so far arbitrary. We shall deal with their magnitudes later. In each of the outer sums in (5.1.2) the numbers $p_1 \ldots p_l$ $(1 \leqslant l \leqslant 2k)$ are only counted once.

Let $d = p_1 p_2 \ldots p_l$, $\mu(d) \neq 0$, $d \equiv 1 \pmod 2$. We evaluate the sum

$$S_d = \sum_{d \mid a_n} 1. \tag{5.1.3}$$

This sum is equal to the number of solutions of the congruence

$$n(n+2) \equiv 0 (\mathrm{mod}\, d), \tag{5.1.4}$$

while the number of solutions of this congruence is equal to the number of solutions $\tau(d)$ of the systems

$$\left.\begin{aligned} n &\equiv 0 (\mathrm{mod}\, d_1), \\ n+2 &\equiv 0 (\mathrm{mod}\, d_2), \end{aligned}\right\} \tag{5.1.5}$$

$d_1 d_2 = d$. But if $n \leqslant x$, the number of solutions of system (5.1.4) is equal to

$$\frac{x}{d} + \theta, \quad |\theta| \leqslant 1$$

and hence, the number of solutions of (5.1.4) is

$$x \frac{\tau(d)}{d} + \theta\tau(d) \qquad (|\theta| \leqslant 1)$$

or

$$S_d = x \frac{\tau(d)}{d} + \theta\tau(d). \tag{5.1.6}$$

If $d \equiv 0 \pmod 2$, then S_d is equal to the number of solutions of the congruence

$$m(m+1) \equiv 0 \left(\mathrm{mod}\, \frac{d}{2}\right) \qquad \left(m \leqslant \frac{x}{2}\right),$$

and we obtain as above:

$$S_d = x \frac{\tau\left(\frac{d}{2}\right)}{d} + \theta\tau\left(\frac{d}{2}\right). \tag{5.1.7}$$

From (5.1.6) and (5.1.7) we can conclude that always

$$S_d = x\frac{\tau'(d)}{d} + \theta\tau'(d), \tag{5.1.8}$$

where

$$\tau'(d) = \begin{cases} \tau(d), & \text{if } d \equiv 1 (\mathrm{mod}\ 2), \\ \tau\left(\frac{d}{2}\right), & \text{if } d \equiv 0 (\mathrm{mod}\ 2). \end{cases} \tag{5.1.9}$$

We substitute (5.1.8) into the right-hand side of (5.1.2):

$$\begin{aligned} \pi_2(x) \leqslant x\Bigg(1 - \sum_{p|R}\frac{\tau'(p)}{p} + \sum_{p_1p_2|R}\frac{\tau'(p_1p_2)}{p_1p_2} \\ - \cdots + \sum_{p_1\cdots p_{2k}|R}\frac{\tau'(p_1\cdots p_{2k})}{p_1\cdots p_{2k}}\Bigg) + \sum_{p|R}\tau'(p) + \sum_{p_1p_2|R}\tau'(p_1p_2) \\ + \cdots + \sum_{p_1\cdots p_{2k}|R}\tau'(p_1\cdots p_{2k}). \end{aligned} \tag{5.1.10}$$

But the sum $\sum_{p_1\cdots p_r|R}\tau'(p_1 \ldots p_r)$ does not exceed

$$2^r C^r_{\pi(z)} < 2^r\frac{\pi^r(z)}{r!}.$$

Consequently,

$$\begin{aligned} \sum_{p|R}\tau'(p) + \sum_{p_1p_2|R}\tau'(p_1p_2) \\ + \cdots + \sum_{p_1\cdots p_{2k}|R}\tau'(p_1\cdots p_{2k}) < \pi^{2k}(z)\sum_{r\leqslant 2k}\frac{2r}{r!} \leqslant 9\pi^{2k}(z). \end{aligned} \tag{5.1.11}$$

To complete our inequality for $\pi_2(x)$ we still have to consider the sum

$$-\sum_{p_1\cdots p_{2k+1}|R}\frac{\tau'(p_1\cdots p_{2k+1})}{p_1\cdots p_{2k+1}} + \sum_{p_1\cdots p_{2k+2}|R}\frac{\tau'(p_1\cdots p_{2k+2})}{p_1\cdots p_{2k+2}} - \cdots = T_{2k}, \tag{5.1.12}$$

which brings up the first term in (5.1.10) to the product

$$\left(1 - \frac{1}{2}\right)\prod_{3\leqslant p\leqslant z}\left(1 - \frac{2}{p}\right). \tag{5.1.13}$$

When considering (5.1.12), we use the inequality

$$\sum_{p_1\cdots p_r|R}\frac{1}{p_1\cdots p_r} \leqslant \frac{\left(\sum_{p\leqslant z}\frac{1}{p}\right)^r}{r!}.$$

It is proved quite easily by induction. But we know that

$$\sum_{p \leqslant z} \frac{1}{p} = \ln \ln z + c.$$

We can therefore write

$$\sum_{p_1 \cdots p_r \mid R} \frac{1}{p_1 \cdots p_r} < \frac{(\ln \ln z + c)^r}{r!}.$$

We conclude from this that the absolute value of (5.1.12) is not greater than the sum

$$\sum_{r \geqslant 2k+1} \frac{(2 \ln \ln z + 2c)^r}{r!}. \tag{5.1.14}$$

We apply to this sum the familiar inequality

$$r! > \left(\frac{r}{e}\right)^r$$

and see that (5.1.14) is not greater than

$$\sum_{r \leqslant 2k+1} \left(\frac{2e \ln \ln z + 2ec}{r}\right)^r.$$

We now put $k = 2e \ln \ln z + 2ec$; we finally find that (5.1.12) does not exceed

$$2^{-2k} < \frac{1}{\ln^4 z}.$$

On taking account of this inequality, we obtain:

$$\begin{aligned} 1 - \sum_{p \mid R} \frac{\tau'(p)}{p} + \sum_{p_1 p_2 \mid R} \frac{\tau'(p_1 p_2)}{p_1 p_2} - \ldots + \sum_{p_1 \cdots p_{2k} \mid R} \frac{\tau'(p_1 \ldots p_{2k})}{p_1 \cdots p_{2k}} \\ = \sum_{d \mid R} \mu(d) \frac{\tau'(d)}{d} - T_{2k} \leqslant \prod_{p \leqslant z} \left(1 - \frac{\tau'(p)}{p}\right) + \frac{1}{\ln^4 z} \\ = \frac{1}{2} \prod_{3 \leqslant p \leqslant z} \left(1 - \frac{2}{p}\right) + \frac{1}{\ln^4 z}. \end{aligned}$$

On substituting this and (5.1.11) in the right-hand side of (5.1.10), we find that

$$\pi_2(x) < \frac{x}{2} \prod_{p=3}^{z} \left(1 - \frac{2}{p}\right) + \frac{x}{\ln^4 z} + 3\pi^{2k}(z). \tag{5.1.15}$$

In order to satisfy

$$\pi^{2k}(z) < \frac{x}{\ln^4 z},$$

it is sufficient to put $z = x^{1/2k}$. If we now use the fact that

$$\prod_{3 \leqslant p < z} \left(1 - \frac{2}{p}\right) < \frac{c_0}{\ln^2 z},$$

where $c_0 > 0$ is an absolute constant, we obtain from (5.1.15):

$$\pi_2(x) < c \frac{x}{\ln^2 x} (\ln \ln x)^2.$$

This last inequality is sufficient for the proof of Theorem 5.1.1.

This sieve has been called "rectangular", because at all the $2k$ steps we have used an invariable number of primes for sieving, equal to $\pi(z)$. In the next section we describe Brun's method, which might be called a "double triangular sieve", since the number of primes used diminishes as the number of steps in the sieve increases.

2. Viggo Brun's Sieve

In 1919 Viggo Brun proposed the following sieve method. Let us be given the arithmetic progression

$$\Delta, \Delta + D, \quad \Delta + 2D, \ldots \qquad ((D, \Delta) = 1, \quad 0 < \Delta < D) \tag{5.2.1}$$

and the sequence of different primes

$$p_1, p_2, \ldots, p_r$$

with two series of their residues, respectively:

$$a_1, \quad a_2, \quad a_3, \ldots, \quad a_r, \quad 0 \leqslant a_i \leqslant p_i.$$
$$b_1, \quad b_2, \quad b_3, \ldots, \quad b_r, \quad 0 \leqslant b_i \leqslant p_i.$$

To simplify the treatment later, we shall assume that $a_i \neq b_i$. We have to investigate the number of numbers of progression (5.2.1) which are $\leqslant x$ and differ from all the numbers of the form

$$\lambda p_i + a_i, \quad \lambda p_i + b_i.$$

Brun denoted this number by the symbol

$$P(\Delta, D, x; a_1 b_1 p_1; a_2 b_2 p_2; \ldots; a_r b_r p_r)$$

or, inasmuch as Brun's formula is independent of Δ, a_i, b_i, by the simpler symbol

$$P(D, x, p_1, p_2, \ldots, p_r).$$

To derive the recurrence formula for this number, we observe that the numbers (5.2.1) falling into the progression

$$\lambda p_r + a_r,$$

again form an arithmetic progression

$$\lambda D p_r + \Delta_r, \quad \Delta_r < p_r D, \tag{5.2.2}$$

where Δ_r is easily determined. The numbers of (5.2.1) which appear in (5.2.2) but are not representable in the form

$$\lambda p_i + a_i, \quad \lambda p_i + b_i$$

with $1 \leqslant i \leqslant r-1$, are

$$P(\Delta_r, Dp_r, x; a_1 b_1 p_1; \ldots; a_{r-1} b_{r-1} p_{r-1}).$$

We carry out a similar argument for the numbers of (5.2.1), representable as

$$\lambda p_r + b_r.$$

As a result, our number (A) is equal to

$$\begin{aligned} &P(\Delta, D, x; a_1 b_1 p_1; \ldots; a_{r-1} b_{r-1} p_{r-1}) \\ &- P(\Delta_z, Dp_z, x; a_1 b_1 p_1; \ldots; a_{r-1} b_{r-1} p_{r-1}) \\ &- P(\Delta'_r, Dp_r, x; a_1 b_1 p_1; \ldots; a_{r-1} b_{r-1} p_{r-1}) \end{aligned}$$

or, more briefly,

$$\begin{aligned} &P(D, x, p_1, p_2, \ldots, p_r) \\ &= P(D, x, p_1, \ldots, p_{r-1}) - (2)\, P(Dp_r, x; p_1, \ldots, p_{r-1}), \end{aligned} \tag{5.2.3}$$

where (2) in the second term symbolizes the sum of two magnitudes of the same type. To simplify the writing, we shall use these symbolic coefficients in future.

We obtain from (5.2.3):

$$\begin{aligned} P(D, x, p_1, \ldots, p_r) &= P(D, x) - (2)\, P(Dp_1, x) \\ &- (2)\, P(Dp_2, x, p_1) - (2)\, P(Dp_3, x; p_1, p_2) - \cdots \\ &- (2)\, P(Dp_r, x; p_1, \ldots, p_{r-1}). \end{aligned} \tag{5.2.4}$$

Formula (5.2.4) is applicable to all the terms of the right-hand side:

$$\begin{aligned} P(D, x, p_1, \ldots, p_r) &= P(D, x) - (2) \sum_{\alpha \leqslant r} P(Dp_\alpha, x) \\ &+ (2^2) \sum_{\alpha=1}^{r} \sum_{\beta<\alpha} P(Dp_\alpha p_\beta, x; p_1, \ldots, p_{\beta-1}). \end{aligned}$$

If we discard some of the terms of the double sum, we obtain, in view of the fact that they are non-negative:

$$P(D, x, p_1, \ldots, p_r) \geqslant P(D, x) - (2) \sum_{\alpha \leqslant r} P(Dp_\alpha, x) + (2^2) \sum_{(\omega)} \sum P(Dp_\alpha p_\beta, x; p_1, \ldots, p_{\beta-1}), \quad (5.2.5)$$

where ω is a domain of summation of α, β which lies wholly in $1 \leqslant \beta < \alpha \leqslant r$.

Now let $r = r_0 > r_1 > r_2 > \ldots > r_n \geqslant 1$ be integers, the precise values of which will be dealt with later.

If we repeat the process of applying formula (5.2.4) and discarding from each even sum certain of its terms, we now get

$$\begin{aligned} P(D, x; p_1, \ldots, p_r) \geqslant{}& P(D, x) - (2) \sum_{\alpha \leqslant r} P(Dp_\alpha, x) \\ &+ (2^2) \sum_{\substack{\alpha \leqslant r \\ \beta > \alpha}} \sum_{\beta \leqslant r_1} P(Dp_\alpha p_\beta, x) \\ &- (2^3) \sum_{\substack{\alpha \leqslant r \\ \alpha > \beta > \gamma}} \sum_{\beta \leqslant r_1} \sum_{\gamma < r_1} P(Dp_\alpha p_\beta p_\gamma, x) \\ &+ (2^4) \sum_{\substack{\alpha \leqslant r \\ \alpha > \beta > \gamma > \delta}} \sum_{\beta \leqslant r_1} \sum_{\gamma < r_1} \sum_{\delta \leqslant r_2} P(Dp_\alpha p_\beta p_\gamma p_\delta, x) \cdots \\ &\cdots + (2^{2n}) \sum_{\substack{\alpha \leqslant r \\ \alpha > \beta > \cdots > \lambda}} \cdots \sum_{\lambda \leqslant r_n} P(Dp_\alpha p_\beta \ldots p_\lambda, x) \\ &- (2^{2n+1}) \sum_{\substack{\alpha \leqslant r \\ \alpha > \beta > \cdots > \lambda > \varrho}} \cdots \sum_{\varrho < r_n} P(Dp_\alpha p_\beta \cdots p_\lambda p_\varrho, x). \end{aligned} \quad (5.2.6)$$

$P(d, x)$ is the number of terms of the arithmetic progression with difference d and initial term $\Delta < d$, not exceeding x. Hence

$$P(D, x) = \frac{x}{d} + \theta, \quad |\theta| \leqslant 1.$$

If each of the $P(d, x)$ in (5.2.6) is replaced by x/d, we obtain an error R not exceeding the number of all the terms having the coefficients 2^k. Hence

$$P(D, x, p_1, \ldots, p_r) > \frac{x}{D}\left(1 - 2 \sum_{\alpha \leqslant r} \frac{1}{p_\alpha} + 2^2 \sum_{\substack{\alpha_1 \leqslant r \\ \alpha_1 > \alpha_2}} \sum_{\alpha_2 \leqslant r} \frac{1}{p_{\alpha_1} p_{\alpha_2}} - \cdots - 2^{2n+1} \sum_{\alpha_1 \leqslant r_0} \cdots \sum_{\alpha_{2n+1} < r_n} \frac{1}{p_{\alpha_1} p_{\alpha_2} \ldots p_{\alpha_{2n+1}}}\right) - R. \quad (5.2.7)$$

Obviously, we obtain all these terms, and only these, among the terms of the product

$$\left(1-\sum_{\alpha_1\leqslant r_0}\frac{2}{p_{\alpha_1}}\right)\left(1-\sum_{\alpha_2\leqslant r_1}\frac{2}{p_{\alpha_2}}\right)\left(1-\sum_{\alpha_3\leqslant r_1}\frac{2}{p_{\alpha_3}}\right)$$
$$\ldots\left(1-\sum_{\alpha_{2n}\leqslant r_n}\frac{2}{p_{\alpha_{2n}}}\right)\left(1-\sum_{\alpha_{2n+1}\leqslant r_n}\frac{2}{p_{\alpha_{2n+1}}}\right), \tag{5.2.8}$$

for which

$$\alpha_1>\alpha_2>\cdots>\alpha_{2n+1}. \tag{5.2.9}$$

The number of them is therefore not greater than

$$Q=(2r_0+1)(2r_1+1)^2\ldots(2r_n+1)^2, \tag{5.2.10}$$

and hence

$$|R|\leqslant Q.$$

Let the sum of the terms of (5.2.8) under the condition (5.2.9) be E. We can now rewrite (5.2.7) as

$$P(D, x, p_1, \ldots, p_r)>\frac{x}{D}E-R.$$

We collect the terms of E separately according to the number of its factors, and let $E^{(i)}$ denote i-tuple summation:

$$E=1-E^{(1)}+E^{(2)}-\cdots+E^{(2n)}-E^{(2n+1)}. \tag{5.2.11}$$

The evaluation of E is carried out in stages. Let E_m be the sum of all the terms of (5.2.8), the indices of which are all $>r_m$ as well as satisfying condition (5.2.9). This is possible only for the first $2m-1$ indices $\alpha_1, \alpha_2, \ldots, \alpha_{2m-1}$, by the conditions of the construction: $\alpha_1\leqslant r^0$, $\alpha_2\leqslant r_1$, $\alpha_3\leqslant r_1$, $\alpha_4\leqslant r_2$, $\alpha_5\leqslant r_2$, $\ldots$, $\alpha_{2m}\leqslant r_m$, $\alpha_{2m+1}\leqslant r_m$. We can again collect them according to the multiplicity of the sums and write

$$E_m=1-E_m^{(1)}+E_m^{(2)}-\cdots+E_m^{(2m-2)}-E_m^{(2m-1)}. \tag{5.2.12}$$

This sum enters into E, and each of its terms $E_m{}^{(i)}$ is part of a term of $E^{(i)}$. For instance,

$$E^{(2)}=2^2\sum_{\substack{\alpha_1\leqslant r_0\\ \alpha>\beta}}\sum_{\alpha_2\leqslant r_1}\frac{1}{p_{\alpha_1}p_{\alpha_2}};$$

$$E_m^{(2)}=2^2\sum_{\substack{\alpha_1\leqslant r_0\\ \alpha_1>r_m,\ \alpha_2>r_m\\ \alpha>\beta}}\sum_{\alpha_2\leqslant r_1}\frac{1}{p_{\alpha_1}p_{\alpha_2}}.$$

In these notations $E = E_{n+1}$.

In order to pass from E_m to E_{m+1}, we introduce elementary symmetric functions of

$$\frac{2}{p_{r_{m+1}+1}}, \quad \frac{2}{p_{r_{m+1}+2}}, \quad \ldots, \quad \frac{2}{p_{r_m}} \ (r_{m+1} < r_m) \tag{5.2.13}$$

and denote them by $S^{(1)}_{m+1}, S^{(2)}_{m+1}, \ldots, S^{(i)}_{m+1}$, so that $i = r_m - r_{m+1}$.

We write the expression, similar to (5.2.12):

$$E_{m+1} = 1 - E^{(1)}_{m+1} + E^{(2)}_{m+1} - \cdots + E^{(2m)}_{m+1} - E^{(2m+1)}_{m+1}.$$

$E^{(i)}_{m+1}$ is built up from terms in which the first i indices $> r_{m+1}$. We now have the following possibilities:

(1) All i indices are not greater than r_m; the such of these terms is equal to $S^{(i)}_{m+1}$.

(2) Only the first index $> r_m$; this yields $E^{(1)}_m S^{(i-1)}_{m+1}$. For, $E^{(1)}_m$ is the sum of all the terms with index $> r_m$, while $S^{(i-1)}_{m+1}$ is the symmetric function embracing all the products with $i - 1$ factors, each index of which $\leqslant r_m$.

(3) Only the two first indices $> r_m$; this gives

$$E^{(2)}_m S^{(i-2)}_{m+1}.$$

$(i + 1)$ All i indices $> r_m$; this gives

$$E^{(i)}_m.$$

All in all we obtain

$$E^{(i)}_{m+1} = S^{(i)}_{m+1} + E^{(1)}_m S^{(i-1)}_{m+1} + E^{(2)}_m S^{(i-2)}_{m+1} + \cdots + E^{(i)}_m, \tag{5.2.14}$$

where the right-hand side breaks off after the $2m$th step, since $\alpha_{2m} \leqslant r_m$ in it. Hence

$$\begin{aligned} E_{m+1} = 1 &- (S^{(1)}_{m+1} + E^{(1)}_m) + (S^{(2)}_{m+1} + E^{(1)}_m S^{(1)}_{m+1} + E^{(2)}_m) \\ &\quad \cdots \cdots \cdots \cdots \\ &- \cdots - (S^{(2m-1)}_{m+1} + E^{(1)}_m S^{(2m-2)}_{m+1} + \cdots + E^{(2m-2)}_m S^{(1)}_{m+1} + E^{(2m-1)}_m) \\ &+ (S^{(2m)}_{m+1} + E^{(1)}_m S^{(2m-2)}_{m+1} + \cdots + E^{(2m-2)}_m S^{(2)}_{m+1} + E^{(2m-1)}_m S^{(1)}_{m+1}) \\ &- (S^{(2m+1)}_{m+1} + E^{(1)}_m S^{(2m)}_{m+1} + \cdots + E^{(2m-2)}_m S^{(3)}_{m+1} + E^{(2m-1)}_m S^{(2)}_{m+1}). \end{aligned}$$

We equate the expression obtained with the products:

$$\begin{aligned} E_m \prod_{\nu=r_{m+1}+1}^{r_m} \left(1 - \frac{2}{p_\nu}\right) &= (1 - E^{(1)}_m + E^{(2)}_m - \cdots + E^{(2m-2)}_m - E^{(2m-1)}_m) \\ &\quad \times (1 - S^{(1)}_{m+1} + S^{(2)}_{m+1} - \cdots \pm S^{(i)}_{m+1}) \\ &= 1 - (S^{(1)}_{m+1} + E^{(1)}_{m+1}) + (S^{(2)}_{m+1} + E^{(1)}_m S^{(1)}_{m+1} + E^{(2)}_m) \\ &\quad \cdots \cdots \cdots \cdots \cdots \cdots \cdots \end{aligned}$$

$$
\begin{aligned}
&- \cdots - (S_{m+1}^{(2m+1)} + E_m^{(1)} S_{m+1}^{(2m)} + \cdots + E_m^{(2m-2)} S_{m+1}^{(3)} \\
&+ E_m^{(2m-1)} S_{m+1}^{(2)}) + (*) + (S_{m+1}^{(2m+2)} + E_m^{(1)} S_{m+1}^{(2m+1)} \\
&+ \cdots + E_m^{(2m-2)} S_{m+1}^{(4)} + E_m^{(2m-1)} S_{m+1}^{(3)}) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&- \cdots + \cdots \pm E_m^{(2m-1)} S_{m+1}^{(i)}.
\end{aligned}
$$

Subsequently we shall so arrange the $S_{m+1}^{(j)}$ that they decrease monotonically and are always < 1 when j increases from 1 to i. Having made this remark, we consider the brackets starting with the sign (*). These brackets decrease in absolute value, since the number of terms in each is equal to $2m$ and they are of alternating sign.

Consequently, the sum of all the brackets which appear in the last product and do not appear in E_{m+1} is less than the first of them.

We introduce the notation:

$$
\pi_m = \prod_{\nu=r_{m-1}+1}^{r_{m-1}} \left(1 - \frac{2}{p_\nu}\right) \qquad (m = 1, 2, \ldots, n),
$$

$$
\pi_{k+1} = \prod_{\nu=1}^{r_k} \left(1 - \frac{2}{p_\nu}\right).
$$

Using this, we can now write the inequality

$$
\begin{aligned}
E_{m+1} &> E_m \pi_{m+1} \\
&- (S_{m+1}^{(2m+2)} + E_m^{(1)} S_{m+1}^{(2m+1)} + \cdots + E_m^{(2m-1)} S_{m+1}^{(3)})
\end{aligned} \tag{5.2.15}
$$

on the assumption that

$$
1 > S_{m+1}^{(1)} > S_{m+1}^{(2)} > \cdots > S_{m+1}^{(i)}.
$$

The following inequalities hold for the elementary symmetric functions $S^{(1)}, S^{(2)}, \ldots, S^{(i)}$ of i distinct positive quantities:

$$
\frac{S^{(1)}}{i} > \frac{2S^{(2)}}{(i-1)\,S^{(1)}} > \frac{3S^{(3)}}{(i-2)\,S^{(2)}} > \cdots > \frac{iS^{(i)}}{1 \,.\, S^{(i-1)}}.
$$

(They can be obtained by introducing the polynomial $f(x) = x^i - S^{(1)}x^{i-1} + S^{(2)}x^{i-2} - \ldots \pm S^{(i)}$ and applying Rolle's theorem to it and its derivatives.) It follows from these inequalities that

$$
S^{(1)} > \frac{S^{(2)}}{S^{(1)}} > \frac{S^{(3)}}{S^{(2)}} > \cdots > \frac{S^{(i)}}{S^{(i-1)}}.
$$

And if $S^{(1)} < 1$, then

$$
1 > S^{(1)} > S^{(2)} > \cdots > S^{(i)}.
$$

In addition, we have

$$S^{(j)} < \frac{(S^{(1)})^j}{j!}, \tag{5.2.16}$$

since this inequality is satisfied for $j = 2$. If it is true for $j - 1$, the inequality

$$\frac{jS^{(j)}}{(i-j+1)\,S^{(j-1)}} < \frac{S^{(1)}}{1}$$

gives us (5.2.16).

Consequently, it remains for us to satisfy the condition

$$S^{(1)} < 1.$$

For this, we have to choose the indices r_m so that the sum of the quantities (5.2.13) is < 1 for each m. We shall choose the primes $p_1, p_2, \ldots, p_r$ in their order of natural increase, with $p_1 \leqslant 3$. Further, let $h_0 > h > 1$ be two positive numbers, where the condition

$$0 < 2 \ln h_0 < 1$$

must be fulfilled for h_0.

In view of the fact that

$$\sum_{3 \leqslant p \leqslant w} \frac{1}{p} = \ln \ln w + c_1 + O\left(\frac{1}{\ln w}\right),$$

$$\prod_{3 < p \leqslant w} \left(1 - \frac{2}{p}\right) = \frac{c^2}{\ln^2 w} + O\left(\frac{1}{\ln^3 w}\right), \tag{5.2.17}$$

there exists $w = w_0$, starting with which

$$0 < \sum_{w < p < w^h} \frac{1}{p} < \ln h_0$$

and

$$\prod_{w \leqslant p \leqslant w^h} \left(1 - \frac{2}{p}\right) > \frac{2}{h_0^2}. \tag{5.2.18}$$

We now choose the r_m as follows: let p_{r_1} be the greatest prime $\leqslant p_r^{1/h}$; p_{r_2} the greatest prime $\leqslant p_r^{1/h^2}$ and so on, and let us continue this process up to p_{r_m}, where $p_{r_m} > w_0$.

Let p_{r_k} be the least prime defined by this method. We obtain, by virtue of (5.2.18):

$$\left.\begin{aligned} 0 < \sigma_1 &= \sum_{\nu=r_1+1}^{r} \frac{1}{p_\nu} < \ln h_0, \\ &\cdots\cdots\cdots\cdots \\ 0 < \sigma_k &= \sum_{\nu=r_k+1}^{r_{k-1}} \frac{1}{p_\nu} < \ln h_0. \end{aligned}\right\} \tag{5.2.19}$$

$$\left.\begin{aligned} \pi_1 &= \prod_{\nu=r_1+1}^{r}\left(1-\frac{2}{p_\nu}\right) > \frac{1}{h_0^2}, \\ &\cdots\cdots\cdots \\ \pi_k &= \prod_{\nu=r_k+1}^{r_{k-1}}\left(1-\frac{2}{p_\nu}\right) > \frac{1}{h_0^2}. \end{aligned}\right\} \tag{5.2.20}$$

Our least prime p_{r_k} must satisfy

$$w_0 < p_{r_k} \leqslant p_r^{1/h^k}.$$

Let $p_{r_{k+1}}$ be the greatest prime $\leqslant w_0$. In this case,

$$0 < \sigma_{k+1} = \sum_{\nu=r_{k+1}+1}^{r_k} \frac{1}{p_\nu} < \ln h_0,$$

$$\pi_{k+1} = \prod_{\nu=r_{k+1}+1}^{r_k}\left(1-\frac{2}{p_\nu}\right) > \frac{1}{h_0^2}.$$

We now try to continue the system of indices independently of p_r. All we need do is to choose them so as to satisfy the inequalities:

$$0 < \sigma_{k+2} = \sum_{\nu=r_{k+2}+1}^{r_{k+1}} \frac{1}{p_\nu} < \ln h_0,$$

$$\cdots\cdots\cdots$$

$$0 < \sigma_{n+1} = \sum_{\nu=1}^{r_n} \frac{1}{p_\nu} < \ln h_0 \,(r_{k+1} = 1),$$

$$\pi_{k+2} = \prod_{\nu=r_{k+2}+1}^{r_{k+1}}\left(1-\frac{2}{p_\nu}\right) > \frac{1}{h_0^2},$$

$$\cdots\cdots\cdots$$

$$\pi_n = \prod_{\nu=r_n+1}^{r_{n-1}}\left(1-\frac{2}{p_\nu}\right) > \frac{1}{h_0^2},$$

$$\pi_{n+1} = \prod_{\nu=1}^{r_n}\left(1-\frac{2}{p_\nu}\right) > \frac{1}{h_0^2}.$$

For these inequalities to be satisfied, it is sufficient to take $p_1 \geqslant 3$ with the condition

$$\frac{1}{p_1} < \ln h_0 \qquad \left(1-\frac{2}{p_1}\right) > \frac{1}{h_0^2} \tag{5.2.21}$$

(since one number can be located in each interval).

Suppose that this process has terminated. The primes $p_{r_{k+1}}$, ..., p_{r_n} depend only on the constants h and h_0, and not on p_r.

It now follows from (5.2.21) that

$$\begin{aligned} R < p_r p_{r_1}^2 p_{r_2}^2 \dots p_{r_k}^2 p_{r_{k+1}}^2 \dots p_{r_n}^2 &\leqslant p_r p_r^{2/h} p_r^{2/h^2} \dots p_r^{2/h^k} p_{r_{k+1}}^2 \dots p_{r_k}^2 \\ &= c(h, h_0)\, p_r^{1+2(1/h+1/h^2+\dots+1/h^k)} \\ &= c(h, h_0)\, p^{h+1-h^{1/k}/h-1} < c' p_r^{h+1/h-1}, \end{aligned} \tag{5.2.22}$$

where c' depends only on h and h_0.

We can now use a recurrence method to obtain from (5.2.15) an asymptotic formula for $E_{n+1} = E$. We introduce the notation:

$$\Phi_{m+1} = S_{m+1}^{(2m+2)} + E_m^{(1)} S_{m+1}^{(2m+1)} + \dots + E_m^{(2m-1)} E_{m+1}^{(3)}, \tag{5.2.23}$$

when (5.2.15) becomes

$$E_2 > E_1 \pi_2 - \Phi_2.$$

But $\pi_2 h_0^2 > 1$, $\pi_1 < 1$, so that

$$E_2 > \pi_2(E_1 - h_0^2 \Phi_2) > \pi_1 \pi_2 (E_1 - h_0^2 \Phi_2).$$

Further, $\pi_2 \pi_3 h_0^4 > 1$,

$$\begin{aligned} E_3 &> \pi_3 E_2 - \Phi_3 > \pi_2 \pi_3 (E_1 - h_0^2 \Phi_2) \\ &\quad - \pi_2 \pi_3 h_0^4 \Phi_3 > \pi_1 \pi_2 \pi_3 (E_1 - h_0^2 \Phi_2 - h_0^4 \Phi_3). \end{aligned}$$

. . . . . .

And finally,

$$E_{n+1} > \pi_1 \pi_2 \dots \pi_{n+1} (E_1 - h_0^2 \Phi_2 - h_0^4 \Phi_4 - \dots - h_0^{2n} \Phi_{n+1}). \tag{5.2.24}$$

Let us consider (5.2.23). We have

$$S_l^{(1)} = 2\sigma_l < 2 \ln h_0 < 1 \qquad (l = 1, 2, \dots, n+1)$$

and

$$S_l^{(j)} < \frac{(S_l^{(1)})^j}{j!}.$$

Putting $\tau = 2 \ln h_0$, we obtain

$$\Phi_{m+1} < \frac{\tau^{2m+2}}{(2m+2)!} + E_m^{(1)} \frac{\tau^{2m+1}}{(2m+1)!} + \dots + E_m^{(2m-1)} \frac{\tau^3}{3!}. \tag{5.2.25}$$

From (5.2.14):

$$\begin{aligned} E_{m+1}^{(i)} &= S_{m+1}^{(i)} + E_m^{(1)} S_{m+1}^{(i-1)} + \dots + E_m^{(i)} \\ &< \frac{\tau^i}{i!} + E_m^{(1)} \frac{\tau^{i-1}}{(i-1)!} + \dots + E_m^{(i)}. \end{aligned} \tag{5.2.26}$$

In particular, $E_{m+1}^{(1)} < \tau + E_m^{(1)}$, and since $E_1^{(1)} = 2\sigma_1 < \tau$, we have

$$E_m^{(1)} < \tau + \tau + \dots + \tau = m\tau. \tag{5.2.27}$$

Using (5.2.25), (5.2.26), (5.2.27), a numerical bound can be found for

$$E_m^{(i)} \quad \text{and} \quad \Phi_m$$

with the aid of a recurrence method. We recall that $E_m^{(i)}$ is the sum of the products of i factors, the indices of which are decreasing and are all $> r_m$. Hence none of the $E_m^{(1)}$, $E_m^{(2)}$, ..., $E_m^{(2m-1)}$ is greater than this quantity.

We shall look for a bound of φ_m with any m. Let m_0 be a fixed index of $E_m^{(i)}$, and let the constants c_0 and k be chosen so that

$$E_{m_0}^{(i)} < c_0(\tau k)^i \tag{5.2.28}$$

($k > 0$ can be chosen arbitrarily, then, having assigned c_0, both constants can be increased arbitrarily). We now obtain from (5.2.26):

$$E_{m_0+1}^{(i)} < \frac{\tau^i}{i!} + c_0(\tau k)^i \left\{ \frac{\left(\frac{1}{k}\right)^{i-1}}{(i-1)!} + \frac{\left(\frac{1}{k}\right)^{i-2}}{(i-2)!} + \cdots + 1 \right\}.$$

If $c_0 \geqslant 1$, the first term can be taken into the bracket without destroying the inequality. Consequently,

$$E_{m_0+1}^{(i)} < c_0(\tau k)^i \, l^{1/k},$$

and we obtain $E_{m_0+2}^{(i)}$ from this by replacing c_0 by $c_0 l^{1/k}$;

$$E_{m_0+2}^{(i)} < c_0 l^{2/k} (\tau k)^i$$

and so on. We obtain in the end:

$$E_{mi}^{(i)} < c_0(\tau k)^i \, l^{m-m_0/k} \qquad (m \geqslant m_0).$$

On substituting this inequality in (5.2.25), we obtain

$$\Phi_{m+1} < \frac{\tau^{2m+2}}{(2m+2)!} + c_0(\tau k)^{2m+2} \, l^{m-m_0/k} \left\{ \frac{\left(\frac{1}{k}\right)^{2m+1}}{(2m+1)!} + \cdots + \frac{\left(\frac{1}{k}\right)^3}{3!} \right\}$$
$$(m \geqslant m_0).$$

Again, we can take the first term inside the bracket without destroying the inequality. We obtain as a result:

$$\Phi_{m+1} < c_0(\tau k)^{2m+2} \, l^{m-m_0/k} \left\{ l^{1/k} - \frac{\left(\frac{1}{k}\right)^2}{2!} - \frac{\frac{1}{k}}{1!} - 1 \right\}. \tag{5.2.29}$$

On substituting this in (5.2.24), we obtain after some transformations:

$$E_{n+1} > \pi_1\pi_2 \dots \pi_{n+1} \times (E_1 - h_0^2\Phi_2 - \dots - h_0^{2m_0}\Phi_{m_0+1} - h_0^{2m_0+2}c_0(\tau k)^{2m_0+4}\, l^{1/k}) \times \left\{l^{1/k} - \frac{1}{2k^2} - \frac{1}{k} - 1\right\} \frac{1 - (l^{1/k}h_0^2(\tau k)^2)^{n-m_0}}{1 - h_0^2(\tau k)^2\, l^{1/k}}. \tag{5.2.30}$$

If $h_0^2(\tau k)^2\, l^{1/k} < 1$, we merely strengthen the incquality by discarding the quantity $[l^{1/k}h_0^2(\tau k)^2]^{n-m_0}$:

$$E_{n+1} > \pi_1\pi_2 \dots \pi_{n+1}\left(E_1 - h_0^2\Phi_2 - \dots - h_0^{2m_0}\Phi_{m_0+1} - \frac{h_0^{2m_0+2}c_0(\tau k)^{2m_0+4}\, l^{1/k}\left(l^{1/k} - \frac{1}{2k^2} - \frac{1}{k} - 1\right)}{1 - l^{1/k}(\tau k)^2 h_0^2}\right). \tag{5.2.31}$$

In order for the value of the bracket to be a maximum, we have to take a k such that $k^2 l^{1/k}$ is a minimum. This $k = \frac{1}{2}$.

We have $E_1^{(1)} < \tau$. If we take $c_0 = 2$, $m_0 = 1$, $i = 1$, then condition (5.2.28) is fulfilled. In addition, recalling that $k = \frac{1}{2}$, $\tau = 2 \ln h_0$, (5.2.31) becomes

$$E = E_{n+1} > \pi_1\pi_2 \dots \pi_{n+1}\left(E_1 - h_0^2\Phi_2 - \frac{h_0^4 2(\ln h_0)^6\, l^2(l^2 - 5)}{1 - (lh_0 \ln h_0)^2}\right) \tag{5.2.32}$$

on condition that (5.2.30), or the equivalent inequality

$$lh_0 \ln h_0 < 1,$$

is fulfilled. For this, it is sufficient to put $h_0 = 1{\cdot}29$, whence

$$\ln h_0 < 0{\cdot}255. \tag{5.2.33}$$

We now have

$$lh_0 \ln h_0 < 0{\cdot}9 < 1.$$

Further, by (5.2.12) and (5.2.19),

$$E_1 = 1 - E_1^{(1)} = 1 - \sum_{\nu=r_1+1}^{r} \frac{2}{p_\nu} = 1 - 2\gamma_1 > 1 - 2\ln h_0.$$

In addition, it follows from (5.2.23) and (5.2.16) that

$$\Phi_2 < \frac{5}{4!}\tau^4 = \frac{5}{4!}(2\ln h_0)^4.$$

On using the numerical values of all the required constants, we easily obtain

$$E > \frac{3}{10}\pi_1\pi_2 \dots \pi_{n+1} = 0{\cdot}3 \prod_{\nu=1}^{r}\left(1-\frac{2}{p_\nu}\right). \qquad (5.2.34)$$

We put $h = 89/69 < 1{\cdot}29 = h_0$. We substitute these values in (5.2.11):

$$P(D, x; p_1, p_2, \dots, p_r) > 0{\cdot}3\,\frac{x}{D}\prod_{\nu=1}^{r}\left(1-\frac{2}{p_\nu}\right) - R,$$

where

$$R < cp_r^{h+1/h-1} = cp_r^{7\cdot9},$$

or

$$P(D, x; p_1, p_2, \dots, p_r) > \frac{x}{D}\,\frac{0{\cdot}3c}{\ln^2 p_r} - c_1 p_r^{7\cdot9}.$$

The conditions

$$\frac{1}{p_1} < \ln h_0 \quad \text{and} \quad 1-\frac{2}{p_1} > \frac{1}{h_0^2}$$

were imposed on the choice of h_0.

To satisfy these conditions, we put $p_1 = 7$. Thus the numbers 2, 3, 5 play no part in the "sieve". In order for them not to appear in our sequence, we put $D = 2 \cdot 3 \cdot 5 = 30$ and consider the progression

$$D\lambda + \Delta, \quad (\Delta, D) = 1, \quad 30\lambda + \Delta \leqslant x.$$

Now let $p_1, p_2, \dots, p_r$ be all the primes (except the first three) $\leqslant x^{1/8}$.

p_r is the greatest prime not exceeding $x^{1/8}$. In this case

$$P(30, x, 7, 11, \dots, p_r) > \frac{6hc\,x}{30(\ln x)^2} - c_1 x^{7\cdot9/8}.$$

As from some $x = x_0$, the right-hand side of this inequality becomes positive and tends to infinity, if $x \to \infty$.

We assumed right at the start that, for each number p_i taking part in the "sieve", two residues a_i and b_i are taken, such that

$$a_i \neq b_i.$$

This condition was imposed in order to simplify the scheme of Brun's method. At the expense of slightly greater complexity, this scheme can also embrace the case when

$$a_i = b_i$$

for certain numbers p_i (the number of which is not greater than say $\ln x$). In this case the factors

$$\left(1-\frac{2}{p_i}\right),$$

corresponding to these numbers and appearing in the product

$$\prod_{j=1}^{r}\left(1-\frac{2}{p_j}\right), \tag{5.2.35}$$

will be replaced by the factors

$$\left(1-\frac{1}{p_i}\right).$$

And if we denote by N_1 the product of the excluded numbers

$$p_{i_1}p_{i_2}\cdots p_{i_q}=N_1,$$

the product

$$\prod_{\substack{p_1\leqslant p_r\\ p_i\mid N_1}}\left(1-\frac{2}{p_i}\right)\prod_{p_i\mid N_1}\left(1-\frac{1}{p_i}\right)$$

appears in inequality (5.2.34) in place of the product (5.2.35). This can be transformed to

$$\prod_{p_i\mid N}\left(1+\frac{1}{p_i-1}\right)\prod_{p_i\leqslant p_r}\left(1-\frac{2}{p_i}\right).$$

Consequently, we obtain in the general case:

$$P(30,x,7,11,\ldots,p_r)>\frac{64\,cx}{30(\ln x)^2}\prod_{p\mid N_1}\left(1-\frac{1}{p-1}\right)-c_1x^{7\cdot 9/8}. \tag{5.2.36}$$

Notice that the inequality

$$P(2,x,3,5,7,\ldots,p_r)\geqslant P(30,x,7,11,\ldots,p_r)$$

is always satisfied. Hence, starting with some $x\geqslant x_0$, we obtain as a completely general lower bound:

$$P(2,x,3,\ldots,p_r)>c\prod_{p\mid N_1}\left(1+\frac{1}{p-1}\right)\frac{x}{\ln^2 x}. \tag{5.2.37}$$

On specifying the residues $a_i=0$, $b_i=-2$, we obtain the following theorem on twin almost prime numbers.

THEOREM 5.2.1. *There exist infinitely many pairs of numbers q, $q+2$, such that each of them has not more than 7 prime factors.*

If we put $x = 2N$ and specify the residues by the table

$$\begin{aligned} a_i &\equiv 0 \pmod{p_i}, \\ b_i &\equiv 2N \pmod{p_i}, \end{aligned} \qquad \text{if} \quad p_i \nmid 2N,$$

and $a_i = b_i = 0$ if $p_i | 2N$, we obtain a second theorem.

THEOREM 5.2.2. *There exists an N_0 such that, with $N > N_0$, there is always a solution of the equation*

$$2N = q_1 + q_2,$$

where q_1 and q_2 are almost prime numbers, having not more than 7 prime factors.

It should be mentioned that the constant 7 is not the best. It can be diminished by the same method, though the working required is much more complicated.

An upper bound for the function

$$P(D, x; p_1, \ldots, p_r)$$

may be obtained by an exactly similar method, the only difference being that we start to reject the "superfluous" numbers in an odd number. We obtain as a result:

$$P(2, x, 3, 5, \ldots, p_r) < c \prod_{p | N_1} \left(1 + \frac{1}{p-1}\right) \frac{x}{\ln^2 x}.$$

It must be mentioned that the constant c obtained in this bound is fairly large.

CHAPTER 6

ATLE SELBERG'S METHOD

1. Selberg's Asymptotic Formulae

In 1946 the Norwegian methematician Atle Selberg[45] proposed a different idea for obtaining upper bounds by the sieve method. These are substantially more accurate than Brun's upper bounds, that we discussed in the last chapter. We shall start our present chapter by explaining the essence of this new idea, taking the example of a binary problem.

We denote by $P(N)$ the number of solutions of the equation

$$N = P_1 + P_2,$$

where N is an even integer, P_1, P_2 are prime and each $> \sqrt{N}$. We denote by Q the product of all primes $\leqslant \sqrt{N}$. In this case, we can use the familiar properties of Möbius functions to write

$$P(N) = \sum_{a_n} \sum_{\substack{d \mid a_n \\ d \mid Q}} \mu(d), \qquad a_n = n(N-n)\,(n \leqslant N). \tag{6.1.1}$$

(In future we shall simplify the writing by taking the condition $d|Q$ for granted and omitting it.) Unfortunately, the strict formula (6.1.1) for $P(N)$ yields nothing in this form.

The idea of Selberg's upper bounds amounts to replacing the sum $\sum_{d \mid a_n} \mu(d)$ in (6.1.1) by another sum $\sum_{\substack{d \mid a_n \\ d \leqslant z}} \varrho_d$, where $z \leqslant N^{1-\varepsilon}$, under the condition that always

$$\sum_{d \mid a_n} \mu(d) \leqslant \sum_{d \mid a_n} \varrho_d, \tag{6.1.2}$$

and achieving the possibility of interchanging the summation.

This idea can be realized as follows in our concrete case. We put:

$$\varrho_1 = \lambda_1 = 1,$$
$$\varrho_d = \sum_{\substack{d=\frac{d_1 d_2}{k},\ k=(d_1,d_2)\\ d_1\leqslant\sqrt{z},\ d_2\leqslant\sqrt{z},\\ \varrho_d=0,\quad d>z.}} \lambda_{d_1}\lambda_{d_2}, \qquad 1 < d \leqslant z, \tag{6.1.3}$$

All the divisors in (6.1.3) are square-free.

For the present the λ_d, $d > 1$, are arbitrary real numbers; we shall consider their magnitudes later.

If conditions (6.1.3) are fulfilled, we obtain

$$\sum_{\substack{d\mid a_n\\ d\leqslant z}} \varrho_d = \left(\sum_{\substack{d\mid a_n\\ d\leqslant\sqrt{z}}} \lambda_d\right)^2 \geqslant \sum_{d\mid a_n} \mu(d).$$

Hence inequality (6.1.2) is actually satisfied, and we have

$$P(N) \leqslant \sum_{a_n} \sum_{\substack{d\mid a_n\\ d\leqslant z}} \varrho_d = \sum_{a_n} \left(\sum_{\substack{d\mid a_n\\ d\leqslant\sqrt{z}}} \lambda_d\right)^2. \tag{6.1.4}$$

If the square of the sum is written as a double sum and the order of summation changed, (6.1.4) becomes

$$P(N) \leqslant \sum_{d_1\leqslant\sqrt{z}} \sum_{d_2\leqslant\sqrt{z}} \lambda_{d_1}\lambda_{d_2} \sum_{\substack{a_n\equiv 0 \left(\mathrm{mod}\ \frac{d_1 d_2}{k}\right)\\ k=(d_1 d_2)}} 1. \tag{6.1.5}$$

The inner sum of the unities is computed like sum (5.1.4) and is equal to

$$N\frac{\tau_N\left(\frac{d_1 d_2}{k}\right)}{\frac{d_1 d_2}{k}} + \theta\tau_N\left(\frac{d_1 d_2}{k}\right), \tag{6.1.6}$$

where

$$\tau_N(d) = \tau\left(\frac{d}{(d, N)}\right). \tag{6.1.7}$$

In addition, we have for the square-free d_1 and d_2:

$$\tau_N\left(\frac{d_1 d_2}{k}\right) = \frac{\tau_N(d_1)\tau_N(d_2)}{\tau_N(k)}. \tag{6.1.8}$$

We introduce the notation

$$f(d) = \frac{d}{\tau_N(d)} \tag{6.1.9}$$

and substitute (6.1.6) in (6.1.5), taking account of (6.1.8) and (6.1.9):

$$P(N) \leqslant N \sum_{d_1 \leqslant \sqrt{z}} \sum_{d_2 \leqslant \sqrt{z}} \frac{\lambda_{d_1}}{f(d_1)} \frac{\lambda_{d_2}}{f(d_2)} f(k) + \sum_{d_1 \leqslant \sqrt{z}} \sum_{d_2 \leqslant \sqrt{z}} |\lambda_{d_1}|\,|\lambda_{d_2}|\,\tau_N(d_1)\,\tau_N(d_2). \tag{6.1.10}$$

In addition, if we introduce the sum

$$R = \sum_{d \leqslant \sqrt{z}} |\lambda_d|\,\tau(d) \tag{6.1.11}$$

and use the fact that $\tau_N(d) \leqslant \tau(d)$, the second double sum in (6.1.10) does not exceed R^2; hence

$$P(N) \leqslant N \sum_{d_1 \leqslant \sqrt{z}} \sum_{d_2 \leqslant \sqrt{z}} \frac{\lambda_{d_1} \cdot \lambda_{d_2}}{f d_1 \cdot f d_2} f(k) + R^2. \tag{6.1.12}$$

Let us investigate the double sum in (6.1.12). We introduce the function

$$f_1(r) = f(r) \prod_{p \mid r} \left(1 - \frac{1}{f(p)}\right). \tag{6.1.13}$$

The equation

$$f(k) = \sum_{d \mid k} f_1(d)$$

holds for this function. Consequently,

$$\begin{aligned} &\sum_{d_1 \leqslant \sqrt{z}} \sum_{d_2 \leqslant \sqrt{z}} \frac{\lambda_{d_1} \cdot \lambda_{d_2}}{f(d_1) \cdot f(d_2)} f(k) \\ &\quad = \sum_{d_1 \leqslant \sqrt{z}} \sum_{d_2 \leqslant \sqrt{z}} \frac{\lambda_{d_1} \cdot \lambda_{d_2}}{f(d_1) \cdot f(d_2)} \sum_{\substack{r \mid d_1 \\ r \mid d_2}} f_1(d) \\ &\quad = \sum_{r \leqslant \sqrt{z}} f_1(r)\, y_r^2, \end{aligned} \tag{6.1.14}$$

where

$$y_r = \sum_{\substack{r \mid d \\ d \leqslant \sqrt{z}}} \frac{\lambda_d}{f(d)}. \tag{6.1.15}$$

We have obtained a quadratic form, and have reached the appropriate point for making use of the arbitrariness of the λ_d, $d > 1$; we are able to choose them so that the value of the quadratic form is a minimum when $\lambda_1 = 1$.

To find this minimum, we replace the condition $\lambda_1 = 1$ by another condition that follows from it. We multiply the right- and left-hand sides of (6.1.15) by $\mu(r)$ and sum over all $r \leqslant \sqrt{z}$.

We obtain as a result:

$$\sum_{r\leqslant\sqrt{z}} \mu(r) y_r = \sum_{r\leqslant\sqrt{z}} \mu(r) \sum_{\substack{r\mid d \\ d\leqslant\sqrt{z}}} \frac{\lambda_d}{f(d)} = \sum_{d\leqslant\sqrt{z}} \frac{\lambda_d}{f(d)} \sum_{r\mid d} \mu(d) = \lambda_1 = 1,$$

i.e.

$$\sum_{r\leqslant\sqrt{z}} \mu(r) y_r = 1. \tag{6.1.16}$$

Thus, we have to find the minimum of the quadratic form (6.1.14) under the auxiliary condition (6.1.16).

We use the following artifice. We consider the new function

$$F(y_1, y_2, \ldots, y_{[\sqrt{z}]}, \omega) = \sum_{r\leqslant\sqrt{z}} f_1(r) y_r^2 - 2\omega \sum_{r\leqslant\sqrt{z}} \mu(r) y_r, \tag{6.1.17}$$

where ω is real but not yet fixed. The minimum of F under condition (6.1.16) is connected with the required minimum by the equation

$$\min E = \min\Big(\sum_{r\leqslant\sqrt{z}} f_1(r) y_r^2\Big) - 2\max\omega. \tag{6.1.18}$$

But we can transform the function F:

$$F(y_1, y_2, \ldots, y_{[\sqrt{z}]}, \omega) = \sum_{r\leqslant\sqrt{z}} f_1(r) \left(y_r - \omega \frac{\mu(r)}{f_1(r)}\right)^2 - \omega_2 \sum_{r\leqslant\sqrt{z}} \frac{\mu_2(r)}{f_1(r)}.$$

We conclude from this that the equation

$$y_r = \max\omega \frac{\mu(r)}{f_1(r)} \tag{6.1.19}$$

is certainly satisfied at the minimum point of F.

Consequently,

$$\min F = -(\max\omega)^2 \sum_{r\leqslant\sqrt{z}} \frac{\mu_2(r)}{f_1(r)}. \tag{6.1.20}$$

To determine $\max \omega$ we substitute (6.1.19) in (6.1.16) and find that

$$\max \omega = \frac{1}{\sum\limits_{r \leqslant \sqrt{z}} \frac{\mu_2(r)}{f_1(r)}}. \tag{6.1.21}$$

On combining equations (6.1.18), (6.1.20) and (6.1.21), we obtain

$$\min \sum_{r \leqslant \sqrt{z}} f_1(r) y_r^2 = \frac{1}{\sum\limits_{r \leqslant \sqrt{z}} \frac{\mu_2(r)}{f_1(r)}}. \tag{6.1.22}$$

Inequality (6.1.12) becomes

$$P(N) \leqslant \frac{N}{\sum\limits_{n \leqslant \sqrt{z}} \frac{\mu_2(r)}{f_1(r)}} + R^2. \tag{6.1.23}$$

To finalize our inequality, we have to compute the sum in the denominator on the right-hand side of (6.1.23) and find an inequality for R.

Let us start with R. We recall that

$$R = \sum_{d \leqslant \sqrt{z}} |\lambda_d| \tau(d).$$

We shall express λ_d in terms of the sum of the values of φ_r with the aid of an inversion formula. We put $r = mk$ in (6.1.15):

$$y_{mk} = \sum_{\substack{mk \mid d \\ d \leqslant \sqrt{z}}} \frac{\lambda_d}{f(d)}.$$

We multiply both sides of this equation by $\mu(k)$ and sum over all $k \leqslant \sqrt{z}/m$, $(k, m) = 1$:

$$\sum_{\substack{k \leqslant \sqrt{z}/m \\ (k,m)=1}} \mu(k) y_{mk} = \sum_{\substack{k \leqslant \sqrt{z}/m \\ (k,m)=1}} \mu(k) \sum_{\substack{mk \mid d \\ d \leqslant \sqrt{z}}} \frac{\lambda_d}{f(d)}$$

$$= \sum_{\substack{k \leqslant \sqrt{z}/m \\ (k,\nu)=1}} \mu(k) \sum_{\substack{k \mid \nu \\ \nu \leqslant \sqrt{z}/m \\ (m,\nu)=1}} \frac{\lambda m\nu}{f(m\nu)} = \sum_{\nu \leqslant \sqrt{z}/m} \frac{\lambda m\nu}{f(m\nu)} \sum_{k \mid \nu} \mu(k) = \frac{\lambda_m}{f(m)}.$$

Consequently,

$$\lambda_m = f(m) \sum_{\substack{k \leqslant \sqrt{z}/m \\ (k,m)=1}} \mu(k) y_{mk}. \tag{6.1.24}$$

We substitute for $\max \omega$ from (6.1.21) in (6.1.19):

$$y_r = \frac{\mu(r)}{f_1(r)} \frac{1}{\sum\limits_{n \leqslant \sqrt{z}} \frac{\mu_2(n)}{f_1(n)}}.$$

Having found y_r, we find λ_m:

$$\lambda_m = \mu(m) \frac{f(m)}{f_1(m)} \frac{1}{\sum\limits_{n \leqslant \sqrt{z}} \frac{\mu^2(n)}{f_1(n)}} \sum_{\substack{k \leqslant \sqrt{z}/m \\ (k,m)=1}} \frac{\mu^2(k)}{f_1(k)}.$$

It is self-evident from this last expression that

$$|\lambda_m| \leqslant \frac{f(m)}{f_1(m)} = \prod_{p \mid m} \left(1 - \frac{1}{f(p)}\right)^{-1}.$$

Let us consider the product. We had $f(r) = 1$, $f(p) \leqslant 2$, if $p \geqslant 3$, hence

$$\prod_{p \mid m} \left(1 - \frac{1}{f(p)}\right)^{-1} \leqslant 2 \prod_{\substack{p \mid m \\ p \geqslant 3}} \left(1 - \frac{2}{p}\right)^{-1}$$

$$\leqslant 2 \prod_{3 \leqslant p \leqslant m} \left(1 - \frac{2}{p}\right)^{-1} < c(\ln m)^2.$$

Replacing the prime divisors of the number m by all the primes up to m obviously leads to a very crude inequality for the product. Nevertheless, this has no effect on the final result. Thus

$$|\lambda_m| \leqslant c(\ln m)^2.$$

On substituting this in R, we obtain

$$R \leqslant c \ln^2 N \sum_{d \leqslant \sqrt{z}} \tau(d) < c \sqrt{z} \ln^3 N,$$

or

$$R^2 < c_1 z \ln^6 N. \tag{6.1.25}$$

To complete the inequality for $P(N)$, we still have to compute

$$S(N, z) = \sum_{n \leqslant \sqrt{z}} \frac{\mu^2(n)}{f_1(n)}.$$

We shall omit the detailed computation, to avoid encumbering the text with long expressions, and merely give the outlines of

the elementary working. The reader can easily fill in the details if he so desires. We split up the sum into factors of the number N:

$$\sum_{n\leqslant\sqrt{z}}\frac{\mu^2(n)}{f_1(n)}=\sum_{\substack{d\mid N\\ d\leqslant\sqrt{z}}}\sum_{\substack{n\leqslant\sqrt{z}\\ (n,N)=1}}\frac{\mu^2(n)}{f(n)\prod\left(1-\frac{1}{f(p)}\right)}$$

$$=\sum_{\substack{d\mid N\\ d\leqslant\sqrt{z}}}\frac{1}{\varphi(d)}\sum_{\substack{n\leqslant\sqrt{z}/d\\ (n,N)=1}}\frac{\mu^2(n)\,\tau(n)}{n\prod_{p\mid n}\left(1-\frac{2}{p}\right)}=\sum_{d\mid N}\frac{S_d}{\varphi(d)}$$

$$=\sum_{\substack{d\mid N\\ d\leqslant\exp(\sqrt{\ln r})}}\frac{S_d}{\varphi(d)}+\sum_{\substack{d\mid N\\ d>\exp(\sqrt{\ln r})}}\frac{S_d}{\varphi(d)}, \tag{6.1.26}$$

where $\varphi(d)$ is Euler's function.

The second step is to obtain an asymptotic formula for S_d with $d\leqslant\exp(\sqrt{\ln r})$:

$$S_d=\sum_{\substack{n\leqslant x\\ (n,N)=1}}\frac{\mu^2(n)\,\tau(n)}{n}\prod_{p\mid n}\left(1+\frac{2}{p-2}\right)$$

$$=\sum_{\substack{n\leqslant x\\ (n,N)=1}}\frac{\mu^2(n)\,\tau(n)}{n}\sum_{d\mid n}\frac{\tau(d)}{\varphi_1(d)}=\sum_{d\mid x}\frac{\mu^2(d)\,\tau^2(d)}{\varphi_1(d)\,d}\sum_{\substack{n\leqslant x/d\\ (n,dN)=1}}\frac{\mu^2(n)\,\tau(n)}{n}.$$

The inner sum is computed by elementary means as follows:

$$\sum_{\substack{n\leqslant x/d\\ (n,Nd)=1}}\frac{\mu^2(n)\,\tau(n)}{n}=\sum_{n\leqslant x/d}\frac{\tau(n)}{n}\left(\sum_{\substack{\nu\mid Nd\\ \nu\mid n}}\mu(\nu)\right)\left(\sum_{r^2\mid n}\mu(r)\right).$$

On changing the order of summation, we reduce everything to the sum

$$\sum_{n\leqslant y}\frac{\tau(n)}{n}=\frac{1}{2}\ln^2 y+O(\ln y).$$

On substituting the last sum in S_d and replacing the sums by products, we find that

$$S_d=\frac{1}{2}\prod_{\substack{p\mid N\\ p\leqslant N}}\left(1-\frac{1}{(p-1)^2}\right)^{-1}\ln^2\frac{\sqrt{z}}{d}+O(\ln z).$$

We obtain from this:

$$\sum_{\substack{d\mid N\\ d\leqslant\exp(\sqrt{\ln z})}}\frac{S_d}{\varphi(d)}=\frac{1}{2}\prod_{p\mid N}\left(1-\frac{1}{p}\right)\prod_{\substack{p\nmid N\\ p\leqslant N}}\left(1-\frac{1}{(p-1)^2}\right)^{-1}\ln^2 z$$
$$+O(\ln^{1.5} z\ln\ln z). \tag{6.1.27}$$

If $d > \exp(\sqrt{\ln z})$, we apply the self-evident formula

$$S_d \leqslant S^1 \leqslant c \ln^2 z.$$

$$\sum_{\substack{d \mid N \\ d > \exp(\sqrt{\ln z})}} \frac{S_d}{\varphi(d)} < c \ln z \sum_{k \geqslant 1} \sum_{\substack{p_1 \cdots p_k \mid N \\ p_1 \cdots p_k > \exp(\sqrt{\ln z})}} \frac{1}{\varphi(p_1 \cdots p_k)}.$$

But the number N can have not more than $2(\ln N)/(\ln \ln N)$ different prime factors.

If $k \leqslant (\sqrt{\ln z})/(\ln \ln^2 N)$, we apply the self-evident formula ($\omega(N)$ is the number of prime divisors of N):

$$\sum_{\substack{p_1 \cdots p_k \mid N \\ p_1 \cdots p_k > \exp(\sqrt{\ln z})}} \frac{1}{\varphi(p_1 \cdots p_k)} \leqslant \omega^k(N) \exp(-\sqrt{\ln z}) \prod_{p \mid N} \frac{1}{\left(1 - \frac{1}{p}\right)}$$

$$\leqslant c\omega^k(N) \ln N \exp(\sqrt{\ln z}) \leqslant \exp\left(-\frac{1}{2}\sqrt{\ln z}\right).$$

Whereas if $k > (\sqrt{\ln z})/(\ln \ln^2 z)$, we have

$$\sum_{\substack{p_1 \cdots p_k \mid N \\ p_1 \cdots p_k > \exp(\sqrt{\ln z})}} \frac{1}{\varphi(p_1 \cdots p_k)} \leqslant \frac{\left(\sum_{p_i \mid N} \frac{1}{p_i - 1}\right)^k}{k!} \leqslant \exp\left(-\frac{1}{2}\sqrt{\ln z}\right).$$

On combining these inequalities, we get

$$\sum_{\substack{d \mid N \\ d > \exp(\sqrt{\ln z})}} \frac{S_d}{\varphi(d)} < c \ln z \exp\left(-\frac{1}{2}\sqrt{\ln z}\right) < c_1.$$

On substituting this inequality and (6.1.27) in (6.1.26), we finally obtain

$$\sum_{n \geqslant \sqrt{z}} \frac{\mu^2(n)}{f_1(n)} = \frac{1}{2} \prod_{p \mid N} \left(1 - \frac{1}{p}\right) \prod_{\substack{p \mid N \\ p \leqslant N}} \left(1 - \frac{1}{(p-1)^2}\right)^{-1} \ln^2 \sqrt{z}$$
$$+ O(\ln^{1.5} z \ln \ln z).$$

We substitute this last formula and (6.1.25) in (6.1.23):

$$P(N) \leqslant 8G(N) \frac{N}{\ln^2 z} + c' \frac{N}{(\ln z)^{2.5}} \ln \ln^2 N + cz \ln^6 N,$$

and if we put $z = N/(\ln^9 N)$, then

$$P(N) \leqslant 8\left(1 + c \frac{\ln \ln^2 N}{\sqrt{\ln N}}\right) G(N) \frac{N}{\ln^2 N},$$

where

$$G(N) = \prod_{p \mid N} \left(1 + \frac{1}{p-1}\right) \prod_{\substack{p \nmid N \\ p < N}} \left(1 - \frac{1}{(p-1)^2}\right).$$

An upper bound can be obtained in precisely the same way for the number of pairs of twin primes:

$$\pi^2(x) \leqslant 8G \frac{x}{\ln^2 x}\left(1 + c\frac{\ln\ln^2 x}{\sqrt{\ln x}}\right),$$

where

$$G = 2 \prod_{N > p > 2} \left(1 - \frac{1}{(p-1)^2}\right).$$

These formulae differ from those assumed by a factor of 8.

2. Schnirelmann's Theorem

In 1930 Schnirelmann[30] published an important application of the upper bound for $P(N)$. By combining this bound with his density method, he obtained the following theorem.

THEOREM 6.2.1. *Every integer is representable as the sum of not more than c_0 primes, where c_0 is an absolute positive constant.*

This theorem represented a first essential step in the solution of additive problems on primes. Schnirelmann obtained $8 \cdot 10^5$ for the value of c_0. This value was subsequently reduced and in 1951 Shapiro and Warg obtained $c_0 \leqslant 20$ by using Selberg's asymptotic formulae. Using more complicated analytic methods, Vinogradov obtained $c_0 \leqslant 4$ for all sufficiently large numbers in 1937.

We turn to the proof of Schnirelmann's Theorem. Let $n' \leqslant x$ be numbers representable as the sums of two primes. We apply Schwartz's inequality:

$$\left(\sum_{n \leqslant x} P(n)\right)^2 \leqslant \left(\sum_{n \leqslant x} P^2(n)\right)\left(\sum_{n' \leqslant x} 1\right).$$

This leads us to a lower bound for the number of numbers $n' \leqslant x$:

$$\sum_{n' \geqslant x} 1 \geqslant \frac{\left(\sum_{n \leqslant x} P(n)\right)^2}{\sum_{n \geqslant x} P^2(n)}. \tag{6.2.1}$$

We have to find a lower bound for the numerator, and an upper bound for the denominator. We take the numerator:

$$\sum_{n\leqslant x} P(n) \geqslant \sum_{\substack{P_1+P_2\leqslant x\\ P_1,P_2\leqslant x/2}} 1 = \pi^2\left(\frac{x}{2}\right) \geqslant 0{,}9\,\frac{x^2}{4\ln^2 x}\,.$$

Consequently,

$$\Big(\sum_{n\leqslant x} P(n)\Big)^2 \geqslant c\,\frac{x^4}{\ln^4 x}\,, \tag{6.2.2}$$

where $c > 0$ is an absolute constant.

As regards the denominator, we make use of an upper bound in the general form:

$$P(n) \leqslant c \prod_{p\mid n}\left(1+\frac{1}{p}\right)\frac{n}{\ln^2 n}\,,$$

$$\sum_{n\leqslant x} P^2(n) \leqslant c_1\,\frac{x^2}{\ln^4 x}\sum_{n\leqslant x}\prod_{p\mid n}\left(1+\frac{1}{p}\right)^2. \tag{6.2.3}$$

We evaluate the sum in (6.2.3):

$$\sum_{n\leqslant x}\prod_{p\mid n}\left(1+\frac{1}{p}\right)^2 = \sum_{n\leqslant x}\Big(\sum_{d\mid n}\frac{1}{d}\Big)^2$$

$$= \sum_{d_1d_2\leqslant x}\sum\frac{1}{d_1d_2}\sum_{\frac{d_1d_2}{k}\mid n,\ n\leqslant x} 1 \leqslant x\sum_{d_1d_2\leqslant x}\sum_{k=(d_1,d_2)}\frac{k}{(d_1d_2)^2} + \sum_{d_1d_2\leqslant x}\sum\frac{1}{d_1d_2}$$

$$\leqslant x\sum_{k\leqslant x}\frac{1}{k^3}\sum_{\substack{\nu_1\nu_2\leqslant x\\ (\nu_1,\nu_2)=1}}\sum\frac{1}{\nu_1^2\nu_2^2} + \sum_{n\leqslant x}\frac{\tau(n)}{n}$$

$$\leqslant x\Big(\sum_{k\leqslant x}\frac{1}{k^3}\Big)\Big(\sum_{\nu\leqslant x}\frac{1}{\nu^2}\Big)^2 + c'\ln^2 x \leqslant c''\,x,$$

where c'' is an absolute constant.

On substituting this inequality in (6.2.3), we obtain

$$\sum_{n\leqslant x} P^2(n) \leqslant c_2\,\frac{x_3}{\ln_4 x}\,. \tag{6.2.4}$$

From inequalities (6.2.4), (6.2.2) and (6.2.1) we find the upper bound:

$$\sum_{n'\leqslant x} 1 > \alpha_0 x, \tag{6.2.5}$$

where $\alpha_0 > 0$ is an absolute positive constant.

We form the sequence of numbers

$$1, n_1', n_2', \ldots, \tag{A}$$

i.e. 1 and all the numbers representable as the sum of two primes. We can say, on the basis of (6.2.5), that the sequence (A) has a positive density in Schnirelmann's sense. Consequently, (A) can serve as a basis for the series of positive integers. Or in other words, there exists an integer k, dependent only on α_0, such that any number m is the sum of not more than k terms of (A):

$$m = \sum_{i \leqslant k} n_i' + \sum 1, \tag{6.2.6}$$

where the sum of the unities has not more than k terms and can be written as a sum of twos and threes, while each $n_i = p_i + p_i'$. Consequently,

$$m = \sum_{\nu \leqslant 3k} p_\nu,$$

and Schnirelmann's Theorem is proved.

CHAPTER 7

THE DISTRIBUTION OF THE FRACTIONAL PARTS OF NUMERICAL SEQUENCES

1. Statement of the Problem

Given the sequence of real numbers

$$\beta_1, \beta_2, \ldots, \beta_n, \ldots \tag{7.1.1}$$

we find the fractional part $\{\beta_j\}$ corresponding to each term β_j (see the definition in [3]) and form the corresponding sequence of fractional parts

$$\{\beta_1\}, \{\beta_2\}, \ldots, \{\beta_n\}, \ldots \tag{7.1.2}$$

All the numbers of (7.1.2) lie in the interval $[0, 1)$. For many problems of the theory of numbers it is important to know the distribution of the numbers of (7.1.2) in $[0, 1)$. We introduce the concept of a distribution function for the fractional parts $\{\beta_j\}$. For $0 \leqslant \gamma < 1$, we denote by $\nu_n(\gamma)$ the number of values of $\{\beta_j\}$ such that $j \leqslant n$ and $\{\beta_j\} < \gamma$. The function

$$V_n(\gamma) = \frac{\nu_n(\gamma)}{n} \tag{7.1.3}$$

indicates the distribution of the first n terms of the sequence (7.1.2) in $[0, 1)$. If the limit

$$\lim_{n\to\infty} V_n(\gamma) = V(\gamma) \tag{7.1.4}$$

exists for each γ, $V(\gamma)$ is called the distribution function for (7.1.2). This function is non-decreasing, and $V(0) = 0$, $V(1) = 1$. The case when $V(\gamma) = \gamma$ is specially important; the sequence (7.1.2) is in this case said to be uniformly distributed (or equi-distributed) over $[0, 1)$. Given a uniformly distributed sequence

(7.1.2), if we specify any closed interval $[\alpha_1, \alpha_2] \subset [0, 1)$, then the number of terms of (7.1.2) lying in $[\alpha_1, \alpha_2]$ is

$$(\alpha_2 - \alpha_1)\, n + o(n), \tag{7.1.5}$$

when n is sufficiently large. It is of the greatest importance to obtain more accurate forms of the remainder term of (7.1.5), i.e. to obtain expressions of the form

$$v_n(\gamma) = \gamma n + O\left(\frac{n}{\varrho(n)}\right), \tag{7.1.6}$$

where $\varrho(n) = n^{\alpha_0}$, $\varrho(n) = (\ln n)^K$ and so on.

In the analytic theory of numbers, problems often arise concerning the distribution of sequences of the form $\{\alpha f(n)\}$, where α is a given number and $f(n)$ is an integral-valued sequence. If α is rational, so that $\alpha = a/q$; $(a, q) = 1$, then $\{\alpha f(n)\}$ can obviously only take the values $0, 1/q, 2/q, \ldots, (q-1)/q$, so that there can be no question of uniform distribution. However, if α is irrational, a uniform distribution of $\{\alpha f(n)\}$ is quite a common phenomenon. Thus, if $f(n) = a_0 n^k + a_1 n^{k-1} + \ldots + a_k$ is an integral-valued polynomial of degree $k \geqslant 1$, then $\{\alpha f(n)\}$ is uniformly distributed. Vinogradov[1] proved in 1937 that the sequence $\{\alpha p\}$, where p runs in succession over the primes and α is irrational, is uniformly distributed. This property of primes lay at the basis of Vinogradov's solution of Goldbach's problem. An elementary proof of the theorem on the uniform distribution of polynomial sequences, together with the relevant asymptotic formulae, was given by Vinogradov.[4] It amounted to an elementary variant of Weyl's method for trigonometric sums.[41] We shall describe the part of this work that we require, together with Vinogradov's work[5] of 1953, in which the sequence $\{\alpha p\}$ is investigated by elementary methods. We shall also discuss an auxiliary theorem on fractional parts.

2. A Lemma on Differences

Following Vinogradov,[4] we shall prove here a lemma which will be required later. Given the system of numbers

$$\beta_1 \leqslant \beta_2 \leqslant \ldots \leqslant \beta_n, \ldots, \tag{7.2.1}$$

lying in the interval $(0, 1]$. We form the fractional parts of the differences

$$\{\beta - \beta'\}, \tag{7.2.2}$$

where β and β' run independently over the numbers $\beta_1, \beta_2, \ldots, \beta_n$. Given $\gamma \in [0, 1)$, let the number T_n of fractional parts (7.2.2) less than γ be given by the formula

$$T_n = n^2\gamma + B\Delta, \tag{7.2.3}$$

where $\Delta \geqslant n$ and B is independent of n. We now have for the sequence (7.2.1), with any $\varrho \in [0, 1)$,

$$\nu_n(\varrho) = n\varrho + B(n\Delta)^{1/3}. \tag{7.2.4}$$

We observe first that each of the β_i can be replaced without loss of generality by the nearest fraction s/n to it on the left, where s is an integer (a new number β_i). To start with, the numbers $\beta - \beta'$ are changed by not more than $2/n$ as a result of this replacement. The number of new numbers $\{\beta - \beta'\}$ lying in the interval $[0, \gamma)$ will, by (7.2.3), differ from the number of old numbers $\{\beta - \beta'\}$ by not more than $n^2(2/n) + B\Delta = 2n + B\Delta = B\Delta$ in view of the inequality for $B\Delta$. The form of (7.2.3) is therefore unchanged. If (7.2.4) is now deduced for the new numbers β_i from (7.2.3), it will follow from (7.2.4) that $B + B\sqrt[3]{n\Delta} = B\sqrt[3]{n\Delta}$ of the new β_i will lie in an interval of length $1/n$, so that (7.2.4) will also hold for the old β_i.

In view of this, we can assume that $\beta_i = s_i/n$, where $0 \leqslant s_i < n$ are integers.

Let the number of numbers (7.2.1) equal to s/n $(0 \leqslant s < n)$ be α_s. Then the number of fractions of series (7.2.2), equal to z/n $(0 \leqslant z < n)$, will be

$$\sum_{s=0}^{n-1} \alpha_s \alpha_{s+z}, \tag{7.2.5}$$

where $s + z$ is defined modulo n in the interval $[0, n)$. We consider the sum

$$\sum_{s=0}^{n-1} \sum_{z=0}^{k-1} \alpha_{s+c}\, \alpha_{s+z}, \tag{7.2.6}$$

where c is any residue modulo n.

By (7.2.3), we have the expression for (7.2.6):

$$kn + B\Delta. \tag{7.2.7}$$

We put

$$S_{s,k} = \sum_{z=0}^{k-1} \alpha_{s+z}.$$

On summing (7.2.6) over $c = 0, 1, \ldots, k-1$, we find that

$$\sum_{s=0}^{n-1} S_{s,k}^2 = k^2 n + Bk\Delta . \tag{7.2.8}$$

We now put

$$S_{s,k} = k + \varepsilon_{s,k} . \tag{7.2.9}$$

Our problem will be to obtain an inequality for $|\varepsilon_{s,k}|$. We obviously have

$$\sum_{s=0}^{n-1} S_{s,k} = kn ; \quad \sum_{s=0}^{n-1} k = kn ,$$

whence

$$\sum_{s=0}^{n-1} \varepsilon_{s,k} = 0 \tag{7.2.10}$$

for any k. Hence, by (7.2.9) and (7.2.8),

$$\sum_{s=0}^{n-1} \varepsilon_{s,k}^2 = \sum_{s=0}^{n-1} (S_{s,k} - k)^2 = \sum_{s=0}^{n-1} S_{s,k}^2 + k^2 n - 2k^2 n = Bk\Delta . \tag{7.2.11}$$

On summing over $k = 1, \ldots, n$, we find that

$$\sum_{k=1}^{n} \sum_{s=0}^{n-1} \varepsilon_{s,k}^2 = Bn^2\Delta . \tag{7.2.12}$$

We put $n\beta_i = s_i$, and represent (see Fig. 1) each s_i as the area of a rectangle bounded by the straight lines $x = i$, $x = i + 1$, $y = 0$, $y = s_i$ $(i = 0, 1, \ldots, n-1)$. Let

$$\max (s_i - i) = \sigma'', \quad \min (s_i - i - 1) = \sigma', \quad \sigma'' - \sigma' = l . \tag{7.2.13}$$

We mark on each horizontal line $y = 0, 1, \ldots, n-1$ its point of intersection with the shaded region (the extreme left-hand point, if there is none). We shall describe these points as marked. Let $M_1(x_1, s')$ and $M_2(x_2, s'')$ be two such points. We obviously have

$$x_2 - x_1 = \alpha_{s_1} + \alpha_{s_1+1} + \cdots + \alpha_{s_2-1} = s_2 - s_1 + \varepsilon_{s_1, s_2 - s_1} ,$$

where s_1, s_2 are the points on the figure corresponding to s', s''. We draw the straight line $y = x - x_1 + s_1$. It cuts the straight line $x = x_2$ at a point $K(x_2, y_2)$. We have

$$y_2 - s_1 = x_2 - x_1 = s_2 - s_1 + \varepsilon_{s_1, s_2 - s_1} = \overline{M_2 K} ,$$

where $\overline{M_2 K}$ is the distance between the points M_2 and K.

The relationship (7.2.12) therefore yields

$$\sum_{M_2,\,K} \overline{M_2K^2} = Bn^2\Delta, \tag{7.2.14}$$

where the summation is over the pairs M_2, K, the same K being determined by the marked points M_1 and M_2. Hence, for the summation of (7.2.14), we need to look over the pairs of marked

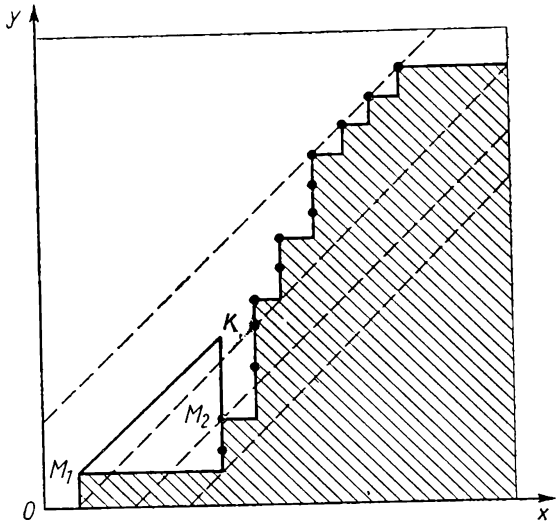

FIG. 1.

points M_1, M_2, find the point K and determine $\overline{MK^2}$. We draw the straight line:

$$y = x + \frac{\sigma' + \sigma''}{2}. \tag{7.2.15}$$

Suppose that the number of marked points above or on this line is $\geqslant \frac{1}{2}n$ (otherwise, this inequality holds for the number of points below or on the line, and the arguments are the same). We draw another straight line:

$$y = x + \frac{3\sigma' + \sigma''}{4}. \tag{7.2.16}$$

Not less than

$$\left[\frac{\sigma'' - \sigma'}{4}\right] \geqslant \frac{l}{4} - 1 \tag{7.2.17}$$

marked points lie below or on it. Each of them can be joinde to each of the points lying on the line (7.2.15) or above it. The corresponding value of $\overline{M_2K}$ is $\geqslant l^2/16$. We have from (7.2.14):

$$\frac{1}{16}l^2\left(\frac{1}{4}l-1\right)\frac{n}{2} = Bn^2\Delta\,, \quad l = B(n\Delta)^{1/3}. \tag{7.2.18}$$

Thus the entire step line lies in the strip between the straight lines $y = x + \sigma'$; $y = x + \sigma''$, where

$$\sigma'' - \sigma' = l = B(n\Delta)^{1/3} \qquad (\sigma'' \geqslant 0;\ \sigma' \leqslant 0).$$

Hence we have: $\varepsilon_0 k = B(n\Delta)^{1/3}$, $S_{s,k} = k + B(n\Delta)^{1/3}$, which proves the lemma.

Successive application of this lemma reduces the problem of the distribution of the sequence $\{P_k(n)\}$, where $P_k(n)$ is a polynomial of degree $k > 1$ with an irrational first coefficient, to the problem of the distribution of the fractional parts of a polynomial of lower degree. It can be shown by this means that, for polynomials $P_k(n)$ of the type mentioned (where k may in fact be equal to 1), $\{P_k(n)\}$ is uniformly distributed. We shall not dwell on this.

3. A Lemma on an Incomplete System of Residues

Let $p > 1$ be a prime, which we shall assume subsequently to be sufficiently large, and let $f(x)$ be an integral-valued function, specified for $x = 0, 1, \ldots, p-1 \pmod p$. Let

$$s = 0, 1, 2, \ldots, p-1\,. \tag{7.3.1}$$

For each s, we form a new integral-valued function with the same domain of specification:

$$f_s(x) = f(x) + sx \tag{7.3.2}$$

and consider the expressions

$$\left\{\frac{f_s(x)}{p}\right\}, \tag{7.3.3}$$

where x runs over the system of numbers $0, 1, 2, \ldots, p-1$. Let $\nu_{fs}(\gamma)$ be the number of fractions (7.3.3) under the condition

$$0 \leqslant \left\{\frac{f_s(x)}{p}\right\} < \gamma \qquad (0 \leqslant \gamma < 1). \tag{7.3.4}$$

Suppose we know that, for small $s = 0, 1, \ldots, p-1$ and any $\gamma \in [0, 1)$,

$$v_{f_s}(\gamma) = \gamma p + Bp^{1-\xi_0}. \tag{7.3.5}$$

Let p_1 be any number $\leqslant p$, and let t run over the (in general, incomplete) system of residues, consisting of the numbers $1, \ldots, p_1$. We are interested in the function $\mu_1(\gamma)$, which indicates how many times the event

$$0 \leqslant \left\{\frac{f(t)}{p}\right\} < \gamma \tag{7.3.6}$$

occurs, when t runs over the above system of residues. By using the method of trigonometric sums developed by Vinogradov, it can be shown without difficulty that

$$\mu_1(\gamma) = \gamma p_1 + Bp^{1-\xi_0} \ln^2 p. \tag{7.3.7}$$

We shall investigate this problem by an elementary method and prove the same relationship. We shall denote the domains of variation of the arguments x and t by (x); (t). We consider the sum

$$\sum_{s=0}^{p-1} \sum_{(x)} \sum_{(t)} V_\gamma\left(\left\{\frac{f(x) + s(x-t)}{p}\right\}\right), \tag{7.3.8}$$

where $V_\gamma(u)$ denotes the function equal to 1 when $0 \leqslant u < \gamma$ and zero when $\gamma \leqslant u < 1$. We shall assume first that γp is not an integer; we shall see from the final result that this involves no loss of generality. When $x = t$ the term

$$p \sum_{(t)} V_\gamma\left(\left\{\frac{f(t)}{p}\right\}\right), \tag{7.3.9}$$

which is the one that interests us, becomes isolated from our sum. Let us consider what terms are obtained when $x \neq t$.

Given x, t, for which $x - t \not\equiv 0 \pmod{p}$, the expression $s(x-t)$ runs over a complete system of residues when $x = 0, 1, \ldots, p-1$; the same applies to the expression $f(x) + s(x-t)$. By the definition of $V_\gamma(u)$, the summation over s in (7.3.8) gives in these conditions

$$[\gamma p], \tag{7.3.10}$$

since γp is assumed not an integer. Hence the sum (7.3.8) is equal to

$$p \sum_{(t)} V_\gamma\left(\left\{\frac{f(t)}{p}\right\}\right) + [\gamma p] p_1 (p-1). \tag{7.3.11}$$

We can consider the sum (7.3.8) from another point of view, by fixing x and s and summing over t. Given s, x, we consider the sum

$$\sum_{(t)} V_\gamma\left(\left\{\frac{f(x)+sx-st}{p}\right\}\right). \tag{7.3.12}$$

When $x = 0$ we obtain the term

$$V_\gamma\left(\left\{\frac{f(x)}{p}\right\}\right)\sum_{(t)} 1.$$

On summing over x, we obtain, in view of condition (7.3.5):

$$\gamma p_1 p + B p_1 p^{1-\xi_0}. \tag{7.3.13}$$

Let $s \neq 0$. We consider the behaviour of the sum (7.3.12). Here t runs over the values $1, 2, \ldots, p_1$. On setting $f(x) + sx = y_s$, we can consider the sum

$$\sum_{(t)} V_\gamma\left(\left\{\frac{y_s - st}{p}\right\}\right). \tag{7.3.14}$$

4. The Comparison of Two Sums

In addition to the sum (7.3.14), we form the sum

$$\sum_{(t)} V_\gamma\left(\left\{\frac{x-st}{p}\right\}\right) \tag{7.4.1}$$

and consider the difference

$$\sum_{(t)} V_\gamma\left(\left\{\frac{y_s - st}{p}\right\}\right) - \sum_{(t)} V_\gamma\left(\left\{\frac{x-st}{p}\right\}\right). \tag{7.4.2}$$

We write the number s/p as

$$\frac{s}{p} = \frac{a_1}{q_1} + \frac{\theta_1}{q_1\tau_1}; \quad (a_1, q_1) = 1; \quad \tau_1 = p_1; \quad |\theta_1| \leqslant 1.$$

If $q_1 \neq 1$, $a_1 \neq 0$. Then $q_1 \leqslant p_1$. We isolate from the interval $[1, p_1]$ of variation of t, starting from the right-hand end, complete systems of residues (mod q_1); let p_1' numbers be thus isolated, and let numbers $1, 2, \ldots, p_2$ remain, forming an incomplete system of residues (mod q_1). We write s/p as

$$\frac{a_2}{q_2} + \frac{\theta_2}{q_2\tau_2}; \quad \tau_2 = p_2; \quad |\theta_2| \leqslant 1; \quad (a_2, q_2) = 1.$$

If $q_2 = 1$, $a_2 = 0$, then we stop here; if not, we continue the process, until we arrive at $q_k = 1$, $a_k = 0$. Since $q_2 \leqslant \frac{1}{2}q_1$; $q_3 \leqslant \frac{1}{2}q_2$, ..., the number of steps in our process will be $B \ln p$. When t runs over a complete system of residues (mod q_1), the quantities $y_s - a_1 t$, $x - a_1 t$ do the same; here, $\theta_1 t/q_1\tau_1 = B/q_1$. Similar arguments can be applied to the moduli $q_2, q_3, \ldots, q_{k-1}$ and on introducing, on the same lines as earlier, the notation $p_1', p_2', p_3', p_4', \ldots$, we obtain the final expression for (7.4.2):

$$R(y_s, s) - R(x, s), \tag{7.4.3}$$

where

$$R(\xi, s) = B\left(\frac{p_1'}{q_1} + \frac{p_2'}{q_2} + \cdots + \frac{p_{k-1}'}{q_{k-1}} + p_k'\right) \tag{7.4.4}$$

for $\xi = 1, 2, \ldots, p-1$. It has been taken into account here that, when $q_k = 1$, $a_k = 0$, we can only give a trivial value for the corresponding sum of the values

$$V_\gamma\left(\left\{\frac{y_s - st}{p}\right\}\right) - V_\gamma\left(\left\{\frac{x - st}{p}\right\}\right).$$

We shall now sum expression (7.4.2) over $x = 0, 1, \ldots, p-1$; $s = 1, 2, \ldots, p-1$. In view of (7.4.3), the sum in question is equal to

$$\sum_{(s)} \sum_{(x)} R(y_s, s) - R(x, s). \tag{7.4.5}$$

We denote by $W_s(y)$ the number of numbers $y_s < y$. Now,

$$\sum_x R(y_s, s) = \sum_x R(x, s)\,(W_s(x+1) - W_s(x)). \tag{7.4.6}$$

We have, by virtue of condition (7.3.5),

$$W_s(y) = y + Bp^{1-\xi_0}. \tag{7.4.7}$$

In view of this and (7.4.4), (7.4.6) takes the form

$$\sum_x R(x, s) + B\left(\frac{p_1'}{q_1} + \frac{p_2'}{q_2} + \cdots + \frac{p_{k-1}'}{q_{k-1}} + p_k'\right) p^{1-\xi_0}, \tag{7.4.8}$$

so that (7.4.5) becomes

$$Bp^{1-\xi_0} \sum_s \left(\frac{p_1'}{q_1}\frac{p_2'}{q_2} + \cdots + \frac{p_{k-1}'}{q_{k-1}} + p_k'\right). \tag{7.4.9}$$

We turn to a consideration of this last sum. We first take the sum

$$\sum_s \frac{p_1'}{q_1} \tag{7.4.10}$$

of the first terms. Here we have: $s/p = a_1/q_1 + \theta/q_1p_1$. The number of numbers s for which q_1 lies between $p_1/2^k$ and $p_1/2^{k+1}$ when $k = 0, 1, 2, \ldots, p_1/2^{k+1} \geqslant \frac{1}{2}$, will have the bound

$$Bp\frac{p_1}{2^k p_1} = B\frac{p}{2^k}.$$

For, it follows from the expression

$$\left|\frac{s}{p} - \frac{a_1}{q_1}\right| \leqslant \frac{1}{q_1 p_1}$$

that

$$|q_1 s - a_1 p| < \frac{p}{p_1};$$

for each given q_1 the number of corresponding values of s will be Bp/p_1; with the indicated values of q_1, the corresponding number of values of s will be $Bp/p_1 \times p_1/2^k = Bp/2^k$.

On multiplying by $p_1/q_1 = B(p_1/p_1)\,2^k = B\,2^k$, and summing over s, we obtain the bound Bp. On summing over the different k, we find the common bound

$$Bp \ln p. \tag{7.4.11}$$

We consider $\sum_s p_2/q_2$ in the same way, taking into account that $\tau_2 = p_2$; $p_2 \leqslant p_2'$, then $\sum_s (p_j'/q_j)$ $(j \leqslant k-1)$. To compute $\sum_s p_k'$, we observe that the number of numbers s for which $s/p = 0/1 + \theta/\tau$, where $\frac{1}{2}p_k \leqslant \tau \leqslant p_k$, will obviously be Bp/τ, so that the corresponding sum has the bound

$$\sum_s p_k' = Bp \lg p. \tag{7.4.12}$$

Finally, we obtain a bound for any value of s and $k = B \lg p$. Hence, substitution of (7.4.11) and (7.4.12) in (7.4.9) gives

$$Bp^{2-\xi_0} \ln^2 p. \tag{7.4.13}$$

On returning to (7.4.2) and summing over (x) and $s = 1, 2, \ldots, p-1$, we find that

$$\sum_{s=1}^{p-1} \sum_{(x)} V_\gamma\left(\left\{\frac{y_s - st}{p}\right\}\right) = \sum_{s=1}^{p-1} \sum_{(x)} V_\gamma\left(\left\{\frac{x - st}{p}\right\}\right) + Bp^{2-\xi_0} \ln^2 p.$$

In the last sum, we perform the summation over x with given s and t, with the obvious result $[\gamma p]$; on then summing over s, t, we obtain

$$[\gamma p]p(p-1) + Bp^{2-\xi_0} \lg^2 p. \tag{7.4.14}$$

Comparison of (7.3.12), (7.3.14) and (7.4.14) gives us

$$p \sum_{(t)} V_\gamma\left(\left\{\frac{f(t)}{p}\right\}\right) = \gamma p_1 p + B p^{2-\xi_0} \ln^2 p, \tag{7.4.15}$$

whence

$$\sum_{(t)} V_\gamma\left(\left\{\frac{f(t)}{p}\right\}\right) = \gamma p_1 + B p^{1-\xi_0} \ln^2 p, \tag{7.4.16}$$

which proves our assertion (7.3.7).

5. Vinogradov's Elementary Derivation of Some Theorems on Sequences of Primes

In 1953 Vinogradov[5] gave an elementary means of investigating the behaviour of the fractional parts $\{\alpha p\}$, where α is a given number and p runs over a sequence of primes. We shall make use of[5] in the present section and the next. We introduce the function

$$\Psi_\gamma(z) = V_\gamma(\{z\}) - \gamma, \tag{7.5.1}$$

where $V_\gamma(u)$ is the function introduced in § 3 ($0 \leqslant \gamma < 1$).

Let $q > 1$ be an integer, $(a, q) = 1$, $0 < a \leqslant q - 1$. An investigation is made of the behaviour of the fractional parts of the system of numbers ap/q, where $p \leqslant N$ ($N \to \infty$) and $q < N$.

THEOREM 7.5.1. (Vinogradov.)

$$\sum_{p \leqslant N} V_\gamma\left(\left\{\frac{ap}{q}\right\}\right) = \gamma\pi(N) + B_\varepsilon N^{1+\varepsilon} F, \tag{7.5.2}$$

where $\pi(N) = \sum_{p \leqslant N} 1$, $\varepsilon > 0$ *is an arbitrarily small constant, and*

$$F = \sqrt{\frac{1}{q} + \frac{q}{N}} + N^{-1/6}. \tag{7.5.3}$$

From the analytic point of view this theorem contains within itself the basic stage of the solution of Goldbach's ternary problem and the derivation of the relevant asymptotic formula. We note that (7.5.2) will be non-trivial, if q increases along with N, more rapidly than N . The proof of Theorem 7.5.1 is based on a number of elementary lemmas.

Given the real number 2, we represent the number α as

$$\alpha = \frac{A}{Q} + \frac{\theta}{Q\tau} \qquad ((A, Q) = 1, \quad 0 < Q \leqslant \tau). \tag{7.5.4}$$

LEMMA 1. *Let β be a real number, c an integer, α the number given by* (7.5.4). *In this case*

$$\left|\sum_{x=c}^{c+Q-1} \Psi_\gamma(\alpha x+\beta)\right| < 2. \tag{7.5.5}$$

We can assume $Q > 2$, since the lemma is trivial when $Q \leqslant 2$. We have:

$$\alpha x+\beta = \frac{Ax+f(x)}{Q}; \quad f(x) = \beta Q + \frac{\theta x}{\tau}.$$

We put: $f(x) = n + \varkappa + \lambda(x)$, where n is an integer ($0 \leqslant \varkappa < 1$, $0 \leqslant \lambda(x) < 1$). We have

$$\{\alpha x+\beta\} = \left\{\frac{r+\varkappa+\varrho(r)}{Q}\right\},$$

where r is the least non-negative residue of $Ax + n$ (mod), and $\varrho(r) = \lambda(x)$. The inequality

$$0 \leqslant \{\alpha x+\beta\} < \gamma$$

can only be satisfied when $r = Q-1;\ 0, 1, \ldots, [\gamma Q]$, where the last value is excluded when γQ is an integer; this gives $< \gamma Q + 2$ values of r. It is certainly satisfied only when $r = 0, 1, \ldots, [\gamma Q - 2]$, i.e. with $> \gamma Q - 2$ values of r. Hence the number of times our inequality is satisfied is of the form $\gamma Q + \Phi$, where $|\Phi| < 2$, so that the sum on the left-hand side of (7.5.5) is equal to

$$(1-\gamma)(\gamma Q+\Phi) - \gamma(Q-\gamma Q-\Phi) = \Phi,$$

which proves our lemma.

LEMMA 2. *Let h be an integer, let x run over X_0 consecutive numbers of the series $1, 2, \ldots, q$; and let y run independently of x over Y_0 numbers of the same series, relatively prime to q. We now have*

$$\sum_x \sum_y \Psi_\gamma\left(\frac{xy+h}{q}\right) < 4q(\ln q)^2. \tag{7.5.6}$$

It is easily computed that, when $q \leqslant 70$, the left-hand side, which does not exceed $q(\varphi)(q)$, is less than the right-hand side. We shall therefore assume $q > 70$. Given y, we represent y/q as

$$\frac{y}{q} = \frac{A_0}{Q_0} + \frac{\theta_0}{Q_0 X_0}, \quad (A_0, Q_0) = 1; \quad 0 < Q_0 \leqslant X_0, \quad |\theta_0| \leqslant 1$$

and so on, until we arrive at some $X_{n+1} = 0$. It may easily be verified by applying Lemma 1 that the part of the double sum

(7.5.6) corresponding to the given y does not exceed in absolute value

$$2\left(\frac{X_0}{Q_0}+\frac{X_1}{Q_1}+\cdots+\frac{X_n}{Q_n}\right),$$

while the absolute value of the whole of the double sum has the upper bound

$$\sum_y\left(\frac{2X_0}{Q_0}+\frac{2X_1}{Q_1}+\cdots+\frac{2X_n}{Q_n}\right). \tag{7.5.7}$$

Here, the numbers $n, Q_0, Q_1, \ldots, Q_n, X_0, X_1, \ldots, X_n$ depend on y. Let us find an upper bound for the sum (7.5.7). We isolate the term $2X/Q$. It corresponds to the system of conditions

$$\frac{y}{q}=\frac{A}{Q}+\frac{\theta}{QX} \qquad ((A,Q)=1;\quad 0<Q\leqslant X),$$

which is equivalent to the system

$$yQ-Aq=t \quad \left((A,Q)=1;\quad 0<Q\leqslant X;\quad |t|<\frac{q}{X}\right). \tag{7.5.8}$$

The last of conditions (7.5.8) yields $X<q/|t|$, so that the term $2X/Q$ can be replaced by $2q/Q|t|$. We replace the third and fourth of conditions (7.5.8) by the cruder conditions

$$0<Q\leqslant q, \quad |t|\leqslant\frac{q}{Q}.$$

When $(Q,q)=\delta$ the first of conditions (7.5.8) is fulfilled only when $|t|=t_1\delta$ (t_1 is an integer). Q and $|t|$ now determine 2δ pairs of values y and A, and the sum of the corresponding terms $2q/Q|t|$ will be $\leqslant 2(q/Qt_1\delta)\,2\delta=4q/Qt_1$.

In view of this, the sum (7.5.7) does not exceed

$$4q\sum_{0<Q\leqslant q}\frac{1}{Q}\sum_{0<t\leqslant q/Q}\frac{1}{t},$$

which is $<4q(\ln q)^2$ when $q>70$, as a simple numerical calculation will show. This proves (7.5.6).

LEMMA 3. *Let x run over X different numbers of the series $1, 2, \ldots, q$, while y runs independently over Y different numbers of the same series, relatively prime to q, $(a,q)=1$. The sum*

$$S=\sum_x\sum_y\Psi_\gamma\left(\frac{axy}{q}\right)$$

is now governed by the inequality

$$|S| < 2(XYq)^{1/2} \ln q. \tag{7.5.9}$$

We apply the Cauchy–Bunyakovskii inequality to the sum S. We get

$$|S|^2 \leqslant X \sum_x \left| \sum_y \Psi_\gamma\left(\frac{axy}{q}\right)\right|^2.$$

In accordance with Vinogradov's fundamental principle for finding the upper bound of a double sum, this inequality will only be strengthened if we carry out the summation over all values $x = \xi = 1, 2, \ldots, q$ in its right-hand side. Let y_1 run over the same values as y, independently of y. We obtain, in view of what has been said,

$$|S|^2 \leqslant X \sum_{\xi=1}^{q} \sum_y \sum_{y_1} \Psi_\gamma\left(\frac{a\xi y}{q}\right) \Psi_\gamma\left(\frac{a\xi y_1}{q}\right) = X \sum_y S_y,$$

where

$$S_y = \sum_{y_1} \sum_{u=0}^{q-1} \Psi_\gamma\left(\frac{u}{q}\right) \Psi_\gamma\left(\frac{uy_1y'}{q}\right) = \sum_u \sum_v \Psi_\gamma\left(\frac{u}{q}\right) \Psi_\gamma\left(\frac{uv}{q}\right),$$

and $y'y \equiv 1 \pmod q$, v is the least non-negative residue of $y_1y' \pmod q$. When y_1 runs over its Y values, v runs over Y different numbers of the series $1, 2, \ldots, q$, relatively prime to q. We split S_y into two sums, the first of which contains the terms with $u < \gamma q$, and the second the terms with $u \geqslant \gamma q$. In each of these sums the factor $\Psi_\gamma(u/q)$ is a constant coefficient of the term $\Psi_\gamma(uv/q)$. We apply Lemma 2 to the last term of each sum separately; we get

$$|S_y| < 4q(\ln q)^2(1-\gamma) + 4q(\ln q)^2\gamma = 4q(\ln q)^2,$$

$$|S|^2 < 4XYq(\ln q)^2; \quad |S| < 2(XYq)^{1/2} \ln q.$$

LEMMA 4. *Let x run over a series of distinct numbers included among X consecutive integers, and y run independently of x over a series of distinct numbers included among Y consecutive integers relatively prime to q. We put*

$$S = \sum_x \sum_y \Psi_\gamma\left(\frac{axy}{q}\right).$$

We now have

$$|S| = BXYF_0; \quad F_0 = \left(\frac{1}{X} + \frac{1}{Y} + \frac{1}{q} + \frac{q}{XY}\right)^{1/2} \ln q. \tag{7.5.10}$$

If $X > q$, we split the interval of variation of x into BX/q intervals, and do the same for the interval of variation of y. After

this, we arrive at the sums considered in Lemma 3. We thus obtain the expressions:

for $X \leqslant q$, $Y \leqslant q$,

$$|S| = B(XYq)^{1/2} \ln q = XY\left(\frac{q}{XY}\right)^{1/2} \ln q = BXYF_0;$$

for $X \leqslant q$, $Y > q$ or for $X > q$, $Y \leqslant q$,

$$|S| = B\frac{Y}{q}(Xqq)^{1/2} \ln q + B\frac{X}{q}(Xqq)^{1/2} \ln q = BXYF_0;$$

for $X > q$, $Y > q$,

$$|S| = B\frac{X}{q}\frac{Y}{q}q^{3/2} \ln q = BXYF_0.$$

LEMMA 5. *Let* $1 \leqslant U < N$, $1 \leqslant \Delta \leqslant U$, $U + \Delta \leqslant N$,

$$S = \sum_x \sum_y \Psi_\gamma\left(\frac{axy}{q}\right),$$

where x, y *run over the integers belonging to two increasing sequences, and* $(y, q) = 1$;

$$U < x \leqslant U + \Delta; \quad \frac{N}{U+\Delta} < y \leqslant \frac{N}{x}, \tag{7.5.11}$$

while in other respects x *and* y *are independent. We now have*

$$|S| = B\frac{N\Delta^2}{u^2}F_1,$$

$$F_1 = \left(\frac{1}{\Delta} + \frac{U^2}{N\Delta} + \frac{1}{q} + \frac{qY^2}{N\Delta^2}\right)^{1/2} (\ln q)^2.$$

We shall assume that $F_1 \leqslant \xi_0$, where, as earlier, $\xi_0 > 0$ is a sufficiently small constant (otherwise the lemma is trivial).

Let r_0 be the greatest integer under the condition $2^{r_0} \leqslant F_1^{-1}$. We isolate from the region (7.5.11) (see Fig. 2) "first", "second", "third", ..., "rth" regions, in accordance with the method illustrated in the figure. The rth region is a rectangle with a base of length $\Delta/2^r$ and a height of the proximate order $N\Delta/U^2\,2^r$ (the left-hand and lower sides of a rectangle are not included in the region). The number of rth regions is 2^{r-1}. In accordance with Lemma 4, the part of the sum S corresponding to one of the rth

regions has the upper bound

$$B\frac{\Delta}{2^r}\frac{N\Delta}{U^2 2^r}\left(\frac{2^r}{\Delta}+\frac{U^2 2^r}{N\Delta}+\frac{1}{q}+\frac{qU^2 2^{2r}}{N\Delta^2}\right)^{1/2}\ln q$$

$$= B\frac{N\Delta^2}{U^2 2^r}F_1(\ln q)^{-1}.$$

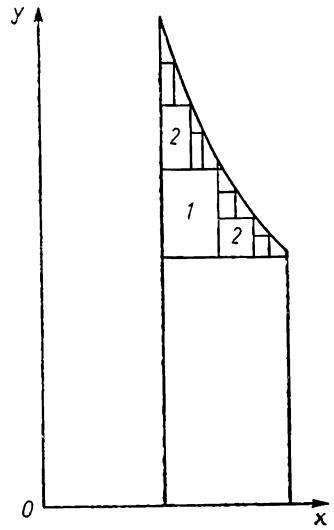

FIG. 2.

The part of S corresponding to the isolated regions has the bound

$$B\frac{N\Delta}{U^2}\frac{\Delta}{2^{r_0}} = B\frac{N\Delta^2}{U^2}F_1.$$

Hence

$$|S| = B\frac{N\Delta^2}{U^2}\left(\sum_{r=1}^{r_0}\frac{2^{r-1}}{2^r\ln q}+1\right)F_1 = B\frac{N\Delta^2}{U^2}F_1,$$

and the lemma is proved.

LEMMA 6. *Let* $1 \leqslant U \leqslant N$, $0 < \Delta \leqslant U$, $U + \Delta \leqslant N$,

$$S_0 = \sum_x \sum_y \Psi_\gamma\left(\frac{axy}{q}\right),$$

$$U < x \leqslant U + \Delta, \quad 0 < y \leqslant \frac{N}{x},$$

while in other respects x, y are independent, and run over the integers of two increasing sequences; $(y, q) = 1$ and not more than M repetitions of each value of y are allowed. We now have

$$|S| = BNMF_2, \quad F_2 = \left(\frac{1}{U} + \frac{U}{N} + \frac{1}{q} + \frac{q}{N}\right)^{1/2} (\ln q)^2.$$

Notice that it is sufficient to consider the case $M = 1$ and $\Delta \geqslant 1$. The part S of the sum S_0, considered in the previous lemma, has the upper bound BNF_2. The remainder of the sum S_0, where the summation is over the region

$$U < x \leqslant U + \Delta, \quad 0 < y \leqslant \frac{N}{U + \Delta},$$

has by Lemma 4 the upper bound

$$B\Delta \frac{N}{U}\left(\frac{1}{\Delta} + \frac{U}{N} + \frac{1}{q} + \frac{qU}{\Delta N}\right)^{1/2} \ln q = BNF_2.$$

LEMMA 7. *Let x, y, m run over the positive numbers relatively prime to q of three increasing sequences. Let*

$$U < x \leqslant U', \quad xym \leqslant N, \quad (x, y) = 1,$$

where

$$1 < U < N, \quad U < U' \leqslant 2U,$$

while in other respects x, y, m are independent. We put

$$S = \sum_x \sum_y \sum_m \Psi_\gamma\left(\frac{axym}{q}\right).$$

We now have

$$|S| = B_\varepsilon N^{1+\varepsilon} F_3, \quad F_3 = \left(\frac{1}{U} + \frac{U}{N} + \frac{1}{q} + \frac{q}{N}\right)^{1/2}.$$

We again assume that $F_3 \leqslant \xi_0$ (ξ_0 is a sufficiently small constant), otherwise the lemma is trivial. We have further:

$$\sum_{d|x;\, d|y} \mu(d) = \begin{cases} 0, & \text{if } (x, y) > 1, \\ 1, & \text{if } (x, y) = 1. \end{cases}$$

In view of this

$$S = \sum_d \mu(d)\, S_d,$$

where

$$S_d = \sum_{x'} \sum_{y'} \sum_m \Psi_\gamma\left(\frac{ad^2x'y'm}{q}\right),$$

and x', y' run over the numbers x/d, y/d, if $(x, y) = d$; the summation is over the region

$$\frac{U}{d} < x' \leqslant \frac{U'}{d}\,; \quad x'y'm \leqslant \frac{N^2}{d}\,.$$

When $N_1 \leqslant N$ the number of pairs m', m under the condition $y'm = N_1$ is equal to $B_\varepsilon N^{\varepsilon 1}$. Hence we have, by Lemma 6, when $d \leqslant F_3^{-1}$,

$$|S_d| = B_{\varepsilon'} \frac{N^{1+\varepsilon'}}{d^2} \left(\frac{d}{U} + \frac{Ud^2}{dN} + \frac{1}{q} + \frac{qd^2}{N} \right)^{1/2} (\ln q)^2$$

$$= B_{\varepsilon'} \frac{N^{1+\varepsilon'}}{d} F_3 (\ln q)^2,$$

while when $d > F_3^{-1}$,

$$|S_d| = B_{\varepsilon'} \frac{N^{1+\varepsilon'}}{d^2} = B_{\varepsilon'} \frac{N^{1+\varepsilon'}}{d} F_3$$

whence the lemma follows.

Though its statement is fairly complicated, Lemma 8 is an elementary lemma from Vinogradov's well-known monograph (see [1]).

LEMMA 8. *Let* $0 < c \leqslant 1/6$, $0 < \sigma \leqslant 1/3$, $0 \leqslant \varrho \leqslant 1 - \sigma$, *and let* P *be a product of primes not exceeding* N^σ; $N > N_0$. *We put* $r = \ln N$,

$$D = r^{\frac{\ln r}{\ln (1+c)}}$$

and consider the divisors d *of* P.

We can distribute them over $D' < D$ *systems, where each system has its own* φ *under the condition that the value of the system satisfies the conditions*

$$\varphi < d \leqslant \varphi^{1+c}.$$

We have $\varphi \leqslant N^\varrho$ *for certain systems, while there exists for each of the rest an integer* $\Gamma > 0$ *and two increasing sequences of positive integers* x, y, *such that* $x \in (\varphi_0, \varphi_0^{1+c}] \subset (N^\varrho, N^{\varrho+\sigma+c}]$. *If we select from all the products* xy *only those for which* $(x, y) = 1$, *we obtain all the numbers* d *of the system in question, each taken* Γ *times, and we obtain only these numbers.*

For the proof, we consider τ, *the greatest integer for each*

$$2^{(1+c)^{\tau-1}} \leqslant N^\sigma.$$

In view of the conditions imposed on c, σ, we have

$$(1+c)^{\tau-1} \leqslant \frac{\ln N}{3 \ln 2}, \quad \tau - 1 \leqslant \frac{\ln \ln N - \ln \ln 8}{\ln(1+c)},$$

$$\tau < \frac{\ln r}{\ln (1+c)} - 1.$$

We put $b = [\ln N]$ and consider all the non-increasing systems b of numbers

$$t_1 \geqslant t_2 \geqslant \cdots \geqslant t_b, \tag{7.5.12}$$

where t_j can take values τ, $\tau - 1$, ..., 1, 0. Let l_s be the number of numbers of the series (7.5.12) equal to s. This series is completely defined by the numbers l_τ, $l_{\tau-1}$, ..., l_1. Hence it follows that the number of such series is less than D.

When $t_j > 0$ we put

$$\varphi_j = 2^{(1+c)^{t_j-1}}, \quad F_j = \varphi_j^{1+c} = 2^{(1+c)^{t_j}}.$$

When $t_j = 0$, we put

$$\varphi_j = F_j = 1.$$

Each $d \leqslant N$ is the product of $\leqslant b$ prime divisors. If this were not the case, we should obtain

$$2 \cdot 3 \cdots (b+2) \leqslant N; \quad \ln 2 + \cdots + \ln (b+2) \leqslant \ln N$$

or

$$\ln N \ln \ln N \leqslant 2 \ln N - 1,$$

which is impossible when $N > N_0$. We arrange the prime divisors of d in decreasing order. If the number of them is $h < b$, we put: $p_{h+1} = \ldots = p_b = 1$. We now obtain: $d = p_1 \ldots p_b$. The series $t_1, \ldots, t_b$ include a unique series under the condition

$$\left.\begin{aligned} \varphi_j < p_j \leqslant F_j \quad &\text{when} \quad t_j > 0, \\ \varphi_j = p_j = F_j \quad &\text{when} \quad t_j = 0. \end{aligned}\right\} \tag{7.5.13}$$

We shall say that the d in question is connected with this series $t_1, \ldots, t_b$. On putting $\varphi = \varphi_1 \ldots \varphi_b$, we obtain: $\varphi < d \leqslant \varphi^{1+c}$. Let $\varphi > N^\varrho$. We consider the system of all d connected with the given series $t_1, \ldots, t_b$. Let β be the least integer for which $\varphi_1 \ldots \varphi_\beta > N^\varrho$ (when $\beta = 1$ we put $\varphi_1 \ldots \varphi_{\beta-1} = 1$).

We have

$$\varphi_1 \cdots \varphi_{\beta-1} \leqslant N^\varrho; \quad \varphi_\beta \leqslant N^\sigma; \quad \varphi_1 \cdots \varphi_\beta \leqslant N^{\varrho+\sigma};$$
$$(\varphi_1 \cdots \varphi_\beta)^{1+c} \leqslant N^{\varrho+\sigma+c}.$$

We put $x = p_1 \dots p_\beta$; $y = p_{\beta+1} \dots p_b$; $\varphi_0 = \varphi_1 \dots \varphi_\beta$. It is easily seen that $x \in (\varphi_0, \varphi^{1+c}] \subset (N^\varrho, N^{\varrho+\sigma+c}]$. Further, let $t_{\beta-k_1+1}, \dots, t_\beta$, $t_{\beta+1}, \dots, t_{\beta+k_2}$ be all the values of t_j equal to t. Let x, y run independently over all the products with the conditions (7.5.13) and

$$p_1 > \dots > p_\beta; \quad p_{\beta+1} \geqslant \dots \geqslant p_b,$$

where the inequality $p_s \geqslant p_{s+1}$ must be understood as the inequality $p_1 > p_{s+1}$ if $p_s > 1$, and as the equality $p_s = p_{s+1}$ if $p_s = 1$. In this case xy coincides with one of the numbers d of our system only if $(x, y) = 1$ and the number of repetitions of the equality $xy = d$ is $C_{k_1+k_2}^{k_1}$. On taking the last number as Γ, we arrive at the proof of Lemma 8. We shall only require a particular case of this lemma, which we shall state as a corollary.

COROLLARY. *Let ν be a constant under the condition $0 < \nu \leqslant 0.1$; let P be a product of primes not exceeding $N^{1/3}$; let d run over the divisors of P not exceeding N;*

$$D = (\ln N)^{\frac{\ln\ln N}{\ln(1+\nu)}}.$$

All the values of d can now be distributed among $D' < D$ classes. Part of these classes includes only values of d under the condition

$$d \leqslant N^{1/3+\nu}.$$

For each of the remaining classes there exists an integer $H > 0$ and two increasing sequences (x) and (y) of positive integers under the condition

$$N^{1/3} < x \leqslant N^{2/3+\nu}$$

such that all the numbers of the class, each taken H times, and only these numbers, are obtained, on selecting among all the products xy only those satisfying the conditions

$$xy \leqslant N \qquad (x, y) = 1.$$

6. Proof of the Theorem

Let $\nu < 0.1$ be a constant, as small as desired. P is the product of all $p \nmid Q$, $p \leqslant N^{1/3}$;

$$Q = \prod_{\substack{N^{1/3} < p \leqslant N \\ p \nmid q}} p,$$

$$S = \sum_{p \leqslant N} \Psi_\gamma\left(\frac{ap}{q}\right). \tag{7.6.1}$$

We have from elementary considerations:

$$\sum_{d \mid p} \sum_{\substack{(m,q)=1 \\ 0<dm\leqslant N}} \mu(d)\, \Psi_\gamma\left(\frac{adm}{q}\right) = S + \frac{1}{2} \sum_{\substack{p_1 \mid Q \\ p_2 \mid Q \\ p_1 p_2 \leqslant N}} \Psi_\gamma\left(\frac{a p_1 p_2}{q}\right) + BN^{2/3}. \tag{7.6.2}$$

Here p_1, p_2 run over the primes and are connected only by the condition $p_1, p_2 \leqslant N$.

All the values $x = p_1$ in the double sum on the right-hand side of (7.6.2) lie in the interval $(N^{1/3}, N^{2/3})$. It can be split into $B \ln N$ intervals as described in Lemma 6. In view of this, application of Lemma 6 gives us the upper bound

$$B_\varepsilon N^{1+\varepsilon}\left(N^{-1/3} + \frac{N^{2/3}}{N} + \frac{1}{q} + \frac{q}{n}\right)^{1/2} = B_\varepsilon N^{1+\varepsilon} F_4$$

for this double sum. We turn to the left-hand side of (7.6.2). Here all the values d can be split into classes, as described in the corollary to Lemma 8. We obtain $B_{\varepsilon'} N^{\varepsilon'}$ classes; $\mu(d)$ is constant for numbers of the same class.

We first consider the class for which we can find, in accordance with the corollary to Lemma 8, an $H > 0$ and two increasing sequences (x), (y) of positive integers under the condition

$$N^{1/3} < x \leqslant N^{2/3+\nu},$$

such that all the numbers of the class, each with H repetitions, and only these numbers, are obtained if we select from all the products xy only those satisfying the conditions

$$xy \leqslant N; \qquad (x, y) = 1.$$

For the chosen class, the corresponding part of the sum on the left-hand side of (7.6.2) becomes, after multiplying by $\mu(d) = \pm 1$ (constants in the class) and the number H,

$$\sum_x \sum_y \sum_m \Psi_\gamma\left(\frac{axym}{q}\right), \tag{7.6.3}$$

where the summation is over the region

$$N^{1/3} < x \leqslant N^{2/3+\nu}; \quad xym \leqslant N; \quad (x, y) = 1.$$

We apply Lemma 7 to the sum (7.6.3). We split the interval of variation of x into $B \ln N$ intervals of the type indicated in Lemma 7. After this, we find the upper bound for (7.6.3):

$$B_\varepsilon N^{1+\varepsilon}\left(N^{-1/3}+\frac{N^{2/3+\nu}}{N}+\frac{1}{q}+\frac{q}{N}\right)^{1/2}=B_{\varepsilon''}N^{1+\varepsilon''}F_4,$$

where $F_4 = F$ (see (7.5.3)).

Further, we must consider the classes mentioned in the corollary to Lemma 8, which contain only values of d subject to $d \leqslant N^{(1/3)+\nu}$. We shall consider only the classes with $\mu(d) = +1$; the case $\mu(d) = -1$ is treated similarly. The part of the sum on the left-hand side of (7.6.2) corresponding to all such classes can be written as $S' + S'' + S'''$, where

$$S'=\sum_{d>N/q}\sum_{m\leqslant N/q}\Psi_\gamma\left(\frac{adm}{q}\right),$$

$$S''=\sum_{d\leqslant N/q}\sum_{m\leqslant[N/qd]q}\Psi_\gamma\left(\frac{adm}{q}\right),$$

$$S'''=\sum_{d\leqslant N/q}\sum_{[N/dq]q<m\leqslant N/d}\Psi_\gamma\left(\frac{adm}{q}\right).$$

The sum S' is non-empty only if $N/q < N^{(1/3)+\nu}$.

In this case we split the interval $[N/q, N^{(1/3)+\nu}]$ of the values $x = d$ into $B \ln N$ intervals of the type mentioned in Lemma 6. On then applying Lemma 6 to S', we obtain the bound

$$B_\varepsilon N^{1+\varepsilon}\left(\frac{1}{Nq^{-1}}+\frac{N^{1/3+\nu}}{N}+\frac{1}{q}+\frac{q}{N}\right)^{1/2}=B_\varepsilon N^{1+\varepsilon}F_4.$$

The part of the sum S'' corresponding to the given d can be written as

$$\left[\frac{N}{dq}\right]\sum_{\substack{0\leqslant\varrho<q\\(\varrho,q)=1}}\Psi_\gamma\left(\frac{\varrho}{q}\right)=\sum_{\delta|q}\mu(\delta)\sum_{u=0}^{q/\delta-1}\Psi_\gamma\left(\frac{u}{q\delta^{-1}}\right).$$

As we know (see [3]), this expression has the upper bound

$$B\tau(q)=B_{\varepsilon'}N^{\varepsilon'},$$

so that

$$|S''|=B\sum_{d<N/q}\frac{N^{1+\varepsilon'}}{dq}=B_\varepsilon\frac{N^{1+\varepsilon}}{q}.$$

The values of d entering into S'' are distributed over the intervals

$$\left[1,\frac{N}{s_0q}\right];\quad\left(\frac{N}{s_0q},\frac{N}{(s_0-1)q}\right];\ldots;\quad\left(\frac{N}{2q},\frac{N}{q}\right],$$

where $s_0 = [\sqrt{N/q}]$. For the part of the sum S''' corresponding to the first interval, we find the upper bound

$$B\frac{N}{s_0 q}q = BN\sqrt{\frac{q}{N}}.$$

We consider the part $S(s)$ of the sum S''' corresponding to the general interval

$$\left(\frac{N}{sq}, \frac{N}{(s-1)q}\right] = I_s.$$

It is equal to

$$\sum_{d\in I_s}\ \sum_{(s-1)q<m\leqslant N/d} \Psi_\gamma\left(\frac{adm}{q}\right).$$

We apply Lemma 5, taking $N/qs(s-1)$ in it instead of Δ and N/sq instead of U. We obtain the bound for $S(s)$:

$$B\frac{N}{s^2}\left(\frac{qs^2}{N}+\frac{1}{q}+\frac{qs^2}{N}\right)^{1/2}(\ln q)^2 = B\frac{N}{s}\left(\frac{1}{q}+\frac{q}{N}\right)^{1/2}(\ln q)^2.$$

Hence

$$\sum_{s=2}^{s_0} S(s) = B_\varepsilon N^{1+\varepsilon}\left(\frac{1}{q}+\frac{q}{N}\right)^{1/2}.$$

On introducing this into (7.6.2), we obtain the formula

$$\sum_{p\leqslant N}\Psi_\gamma\left(\frac{ap}{q}\right) = B_\varepsilon N^{1+\varepsilon}F, \tag{7.6.4}$$

whence Theorem 7.5.1 follows at once.

CHAPTER 8

COMPUTATION OF THE INTEGRAL POINTS WITHIN CONTOURS

1. Statement of the Problem. Typical Problems†

Given a plane, smooth,† closed curve C (see Fig. 3) referred to Cartesian coordinates xOy, and expressed by the equations $y = \varphi_1(x)$ (the lower part) and $y = \varphi_2(x)$ (the upper part). We shall assume that the smoothness of the curve is such that the first and second derivatives $\varphi_i(x)$, $\varphi''(x)$ exist inside the interval $[P, Q]$ (this may not be true at the points P and Q).

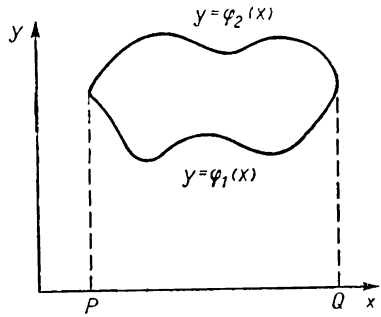

FIG. 3.

The area enclosed by the curve is given by

$$S = \int_P^Q (\varphi_2(x) - \varphi_1(x))\, dx. \tag{8.1.1}$$

† The present chapter makes systematic use of the concept of a smooth curve, so that, to the extent indicated in the foreword, it uses the elements of analysis.

We take in the same coordinate system a lattice of integral points, i.e. a system of points (x, y) with integral coordinates, and mark the points lying inside and on the contour. The number of these points can obviously be written as

$$\sum_{P \leqslant x \leqslant Q} [\varphi_2(x)] - \sum_{P \leqslant x \leqslant Q} [\varphi_1(x)], \tag{8.1.2}$$

where the integral points of the contour, with abscissae P or Q, if they exist, are not taken into account (they can be added to (8.1.2) if required). It is natural to equate (8.1.2) and (8.1.1) if the area S is sufficiently large and the contour is not too "twisting". We shall do this in the cases when $Q - P$ is large and the curvature of the contour is sufficiently small (the radius of curvature is sufficiently large). If L is the perimeter of the curve (the length of the contour), the number of integral points inside and on the contour differs from the area S enclosed by the curve by not more than the number of horizontal strips, bounded by $y = y_1$, $y = y_1 + 1$, or vertical strips, bounded by $x = x_1$, $x = x_1 + 1$, intersected by the contour. The number of such strips does not exceed BL. If N is the number of integral points inside and on the contour, we thus have

$$N = S + BL. \tag{8.1.3}$$

The behaviour of the remainder term $N - S = R$ depends on the curve C. Historically, the first curves investigated in this context were circles and hyperbolas together with their asymptotes.

In the case of a circle $x^2 + y^2 = R^2$ the perimeter $L = 2\pi R$ and the enclosed area $S = \pi R^2$, so that we have

$$N = \pi R^2 + \varrho(R); \quad \varrho(R) = BR. \tag{8.1.4}$$

For the hyperbola $xy = n$ (n is an integer) together with its asymptotes (the coordinate axes $x = 0$ and $y = 0$), the problem of computing the integral points inside the contour was posed by Dirichlet in connection with the problem of divisors. Let $\tau(m) = \sum_{\delta \mid m} 1$ be the number of divisors of the number m; let $T(n) = (\tau(1) + \tau(2) + \ldots + \tau(n))/n$ be the mean number of divisors for numbers $m \leqslant n$. We can write

$$nT(n) = \sum_{m \leqslant n} \tau(m) = \sum_{xy \leqslant n} 1 = \sum_{x \leqslant n} \sum_{y \leqslant n/x} 1 = \sum_{x=1}^{n} \left[\frac{n}{x}\right]. \tag{8.1.5}$$

Obviously, (8.1.5) is the number of integral points in the region $xy \leqslant n$ $(x > 0, y > 0)$. Here, we must obviously have $x \geqslant 1$, $y \geqslant 1$, whence $x \leqslant n$, $y \leqslant n$. We obtain the region illustrated in Fig. 4. The length of the contour L is Bn, so that we have, by (8.1.3),

$$N = S + Bn. \tag{8.1.6}$$

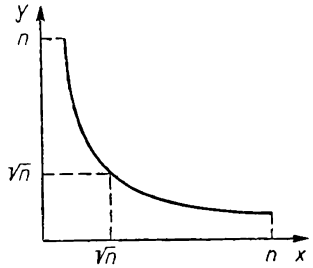

FIG. 4.

Further, since $y = n/x$ is a decreasing function, we have

$$S = n \sum_{m \leqslant n} \frac{1}{m} + Bn, \tag{8.1.7}$$

as is easily seen from the figure. On returning to (8.1.5), we get

$$T(n) = \sum_{m \leqslant n} \frac{1}{m} + B. \tag{8.1.8}$$

We could have obtained the same formula directly from (8.1.5), by observing that $[n/x] = n/x + \{n/x\}$; since $0 \leqslant \{n/x\} < 1$, (8.1.8) follows from (8.1.5). However, the geometrical approach to computing the number of integral points in a contour permits of an almost immediate improvement in the estimate of the remainder term $N - S$. In fact, in view of the symmetry of the contour with respect to the bisector $y = x$, we can see directly from the figure that

$$N = 2N_1 - [\sqrt{n}]^2,$$

where N_1 is the number of integral points in the left-hand region cut out from the given region by the vertical line $x = \sqrt{n}$. Hence

$$\begin{aligned} N = nT(n) &= 2 \sum_{x \leqslant \sqrt{n}} \left[\frac{n}{x}\right] - [\sqrt{n}]^2 \\ &= 2 \sum_{x=1}^{\sqrt{n}} \frac{n}{x} - n + B\sqrt{n} = S + B\sqrt{n}, \end{aligned}$$

in accordance with the same reasoning by which we derived (8.1.7).

Hence, on putting

$$V = 2\sum_{x=1}^{\sqrt{n}} \frac{n}{x} - n,$$

we have

$$N = V + \varrho(n); \quad \varrho(n) = B\sqrt{n}. \tag{8.1.9}$$

From the beginning of this century, the investigations of many mathematicians, starting with G. F. Voronoi[6] and I. Sierpinski[49] have been directed to the problem of finding "strict estimates" of the error $\varrho(R)$ in formula (8.1.4) and $\varrho(n)$ in formula (8.1.9). We shall explain what this means.

Let $R \to \infty$ in (8.1.4). We have: $\varrho(R) = O(R)$. There can exist and (as we shall see) there in fact do exist positive α such that $\alpha < 1$, and $\varrho(R) = O(R^{\alpha})$. We denote the lower bound of such numbers by α_0. We now have, for any $\varepsilon > 0$,

$$\varrho(R) = B_{\varepsilon} R^{\alpha_0+\varepsilon} \quad \text{(for all } R\text{)}.$$

At the same time, the relationship

$$\varrho(R) = B_{\varepsilon} R^{\alpha_0-\varepsilon} \quad \text{(for all } R\text{)}$$

is false, whatever the $\varepsilon > 0$.

The number α_0 is in fact called the strict order of the error in (8.1.4). The strict order of the error in (8.1.9) is similarly defined. The problem of finding the strict order is called the *circle problem* in the former case, and the *divisor problem* in the latter.

In 1914 Hardy[39] proved that, for the circle problem, $\alpha_0 \geqslant \frac{1}{2}$. It was shown by Jarnik[40] that $\alpha_0 < 2/3$. But both the circle and the divisor problems essentially remain unsolved even today.

In 1917 Vinogradov[1] proposed an elementary method for computing the integral points in contours of a general type, which will be described below. He gave $\alpha_0 \leqslant 2/3$ for the circle problem, and $\alpha_0 \leqslant 2/3$ for the divisor problem.

2. The Statement of Vinogradov's Theorem

We shall consider the number of integral points inside and on the trapezoid with the curvilinear top side $y = f(x)$, where $y' \geqslant 0$ is a non-decreasing function, so that $y'' > 0$; let $\alpha \leqslant y'' < \beta$ from P to Q (Fig. 5). If we magnify the curve similarly, the length PQ

and all the ordinates are magnified the same number n of times; y' is unchanged, while y'' is diminished n times and is replaced by y''/n. In view of this, on similar magnification of the contour n times, $\alpha/n < y'' < \beta/n$. But apart from the similitude transformation, there can obviously be other more general variations of the curve, distending it to infinity, such that, when PQ is magnified n times, we still have $\alpha/n < y'' < \beta/n$. It turns out that this latter

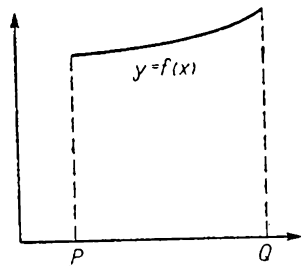

FIG. 5.

condition is in fact the most important for estimating the remainder term, in the sense that, if the curve is distended while preserving this condition, the estimate of the remainder term still remains sufficiently accurate.

We have to seek

$$\sum_x [f(x)] = \sum_x f(x) - \sum_x \{f(x)\}.$$

The main problem, of finding $\sum_x \{f(x)\}$, is arithmetical.

Let the form of $f(x)$ and the two numbers Q and R be determined by a parameter A which $\to \infty$; and let $R - Q \to \infty$, as well as $A \to \infty$. We require an asymptotic formula for the sum

$$S = \sum_{x>Q}^{x \leqslant R} \{f(x)\} \quad \text{as} \quad A \to \infty.$$

We shall impose a restriction on $f(x)$ by assuming that $f'(x)$ exists and is monotonically increasing, and $f''(x) > 0$, where $1/kA \leqslant f''(x) \leqslant 1/A$ for all $A > A_0$; k is fixed and $\geqslant 1$; x varies from Q to R; $Q \leqslant x \leqslant R$.

As A increases, let $R - Q$ increase so that

$$\lim_{A\to\infty} \frac{(A \ln A)^{2/3}}{R - Q} = 0. \tag{8.2.1}$$

THEOREM 8.2.1 (Vinogradov.) *Let k, A, Q, R be numbers under the conditions: $k \geqslant 1$, $A > 25$, $Q < R$; for $Q \leqslant x \leqslant R$, $f(x)$ has a second derivative $f''(x)$, where $1/kA \leqslant f''(x) \leqslant 1/A$. We now have*

$$\sum_{x>Q}^{x\leqslant R} \{f(x)\} = \frac{R-Q}{2} + G, \tag{8.2.2}$$

where

$$|G| \leqslant 2k\left(\frac{R-Q}{A} + 1\right)(A \ln A)^{2/3}. \tag{8.2.3}$$

Before proving the theorem, we require several preliminary lemmas.

LEMMA 1. *If a is a real number, given any $\tau > 1$ we can satisfy the inequalities*

$$-\frac{1}{\tau} < aX - Y < \frac{1}{\tau}, \quad 0 < X < \tau,$$

i.e.

$$\left|a - \frac{Y}{X}\right| < \frac{1}{X\tau}, \quad X < \tau,$$

by the relatively prime numbers X and Y.

This is a familiar theorem of the theory of continued fractions (see [29]).

LEMMA 2. *Given the n positive integers*

$$X_1, X_2, \ldots, X_n; \quad 0 < X_i \leqslant N (i = 1, 2, \ldots, n). \tag{8.2.4}$$

Without destroying the consecutive arrangement of more than N of them, we can group all the remainder in such a way that the number m of terms

$$X_{\alpha+1}, X_{\alpha+2}, \ldots, X_{\alpha+m}$$

of each group is equal to the greatest of the participating numbers, so that

$$X_{\alpha+i} \leqslant m \qquad (i = 1, 2, \ldots, m).$$

The lemma is trivial when $n < N$.

Let $n \geqslant N$. We consider the group of the X_1 first terms of (8.2.4): $X_1, X_2, \ldots, X_{X_1}$. Let $X^{(2)}$ be the greatest of these numbers. Now, either $X^{(2)} = X_1$, or $X^{(2)} > X_1$. If $X^{(2)} = X_1$, the group is isolated. If $X^{(2)} > X_1$, then we take the group of the $X^{(2)}$ first terms of (8.2.4): $X_1, X_2, \ldots, X_{X^{(2)}}$. Let $X^{(3)}$ be the greatest of these numbers. Either $X^{(3)} > X^{(2)}$, or $X^{(3)} = X^{(2)}$.

If $X^{(3)} = X^{(2)}$, the isolation is complete. If not, $X^{(3)} > X^{(2)}$, and we take the group consisting of $X_1, X_2, \ldots, X_{X^{(3)}}$. We repeat the same process. The number of terms of the groups obtained is always increasing: $X_1 < X^{(2)} < X^{(3)} < \ldots$, but all the terms are $\leqslant N$. Thus there are not more than N terms here, i.e. the isolation can be completed by considering not more than N groups. If, after the first isolation, there were to remain $< N$ terms, the lemma would be proved; otherwise, we start the isolation again with the first remaining term, and so on.

Note. Let $X_{\alpha+1}, X_{\alpha+2}, \ldots, X_{\alpha+m}$ be one of the groups. Then $X_i \leqslant m$ $(i = \alpha + 1, \alpha + 2, \ldots, \alpha + m)$. Hence

$$1 \leqslant \sum_{i=\alpha+1}^{i=\alpha+m} \frac{1}{X_i}.$$

On adding for all the groups, we find that the number of groups does not exceed

$$\sum_{i=1}^{n} \frac{1}{X_i}, \quad \text{since the number of groups} = s1 \leqslant \sum_{i=1}^{n} \frac{1}{X_i}.$$

LEMMA 3. *Let $X > 0$ be an integer; Y an integer relatively prime to X; $(Y, X) = 1$; M an integer and $P(z)$ a function of z such that, for any values of α and β from the series $z = M + 1, M + 2, \ldots, M + X$, we have $|P(\alpha) - P(\beta)| < C$, where C is a given number. We now have*

$$-C - \frac{1}{2} \leqslant \sum_{z=M+1}^{M+X} \left\{\frac{Yz + P(z)}{X}\right\} - \frac{1}{2} X \leqslant C + \frac{1}{2},$$

or

$$\left| \sum_{z=M+1}^{M+X} \left\{\frac{Yz + P(z)}{X}\right\} - \frac{X}{2} \right| \leqslant C + \frac{1}{2}.$$

We use the notation

$$S = \sum_{z=M+1}^{M+X} \left\{\frac{Yz + P(z)}{X}\right\}$$

and distinguish two cases:

(1) $C + \frac{1}{2} \geqslant X/2$. But obviously, $0 \leqslant S < X$, so that $-X/2 \leqslant S - X/2 < X/2$ or $-C - \frac{1}{2} \leqslant S - X/2 < C + \frac{1}{2}$, which proves the lemma.

(2) $C + \frac{1}{2} < X/2$.

Let P be the least value of $P(z)$ with $z = M + 1, M + 2, \ldots, M + X$ and $P(z) = P + \Phi(z)$, where $0 \leqslant \Phi(z) \leqslant C$ for all the values of z in question.

Let $[P] = K$; $\{P\} = \varepsilon$. We can now write S as

$$S = \sum_{z=M+1}^{M+X} \left\{\frac{Yz + K + \varepsilon + \Phi(z)}{X}\right\}.$$

We can replace $Yz + K$ by its least residue modulo X. Since $(X, Y) = 1$, u will run in a certain order over the numbers

$$0, 1, 2, \ldots, X - 1,$$

when $z = M + 1, \ldots, M + X$. The sum S takes the form

$$S = \sum_{u=0}^{u=X-1} \left\{\frac{u + \varepsilon + \psi(u)}{X}\right\},$$

where $0 \leqslant \psi(u) \leqslant C$ for all $u = 0, 1, 2, \ldots, X - 1$. We divide all the terms of our series into two groups. The numbers under the condition $0 \leqslant u < X - C - \varepsilon$ are in the first group, and under $X - C - \varepsilon \leqslant u < X$ in the second. We have $0 \leqslant (u + \varepsilon + \psi(u))/X < 1$ for numbers of the first group, and $0 \leqslant (u + \varepsilon + \psi(u))/X < 2$ for numbers of the second, since $X/2 > C + \frac{1}{2}$. Thus, if we put

$$\left\{\frac{u + \varepsilon + \psi(u)}{X}\right\} = \frac{u + \varepsilon + \psi(u)}{X} + h(u),$$

we have $h(u) = 0$ for numbers of the first group, and $h(u) = 0$ or $h(u) = -1$ for numbers of the second. In the second group there are not more than $C + \varepsilon$ numbers

$$\sum_{u=0}^{u=X-1} \frac{u + \varepsilon + \psi(u)}{X} - C - \varepsilon \leqslant S \leqslant \sum_{u=0}^{u=X-1} \frac{u + \varepsilon + \psi(u)}{X},$$

$$\frac{X(X-1)}{2}\frac{1}{X} - C - \varepsilon + \varepsilon\frac{X}{X} \leqslant S, \quad \text{or} \quad \frac{X}{2} - \frac{1}{2} - C \leqslant S.$$

Similarly, $S \leqslant \frac{1}{2}X - \frac{1}{2} + \varepsilon + C$, so that, since $-\frac{1}{2} + \varepsilon < \frac{1}{2}$,

$$\frac{1}{2}X - \frac{1}{2} - C \leqslant S \leqslant \frac{1}{2}X + \frac{1}{2} + C,$$

$$-\frac{1}{2} - C \leqslant S - \frac{1}{2}X \leqslant \frac{1}{2} + C,$$

which it was required to prove.

We turn to the proof of the theorem, which will be performed in several stages.

(1) Let l and $l + 1$ lie in the interval (Q, R). We now have

$$f'(l + 1) - f'(l) = f''(l + \delta)\ (0 < \delta < 1).$$

Hence, in the series

$$f'([Q]+1);\quad f'([Q]+2),\ \ldots,\ f'([R]) \tag{8.2.5}$$

the difference between two adjacent numbers is $\geqslant 1/kA$ and $\leqslant 1/A$. If the least number of (8.2.5) is denoted by K, the greatest of them is $\leqslant K + (R-Q)/A$. Further, the number of numbers of series (7.2.5) not departing outside the limits U and $V > U$ will be $\leqslant kA(U-V)+1$.

(2) We take a number τ satisfying $4 < \tau < \sqrt{A}$. By Lemma 1, there exist one or more pairs X, Y under the conditions $-1/\tau < Xf'(x) - Y < 1/\tau$; $(X, Y) = 1, 0 < X < \tau$. We select one pair, denoting it by $X(x)$, $Y(x)$. Putting $x = [Q]+1, [Q]+2, \ldots, [R]$, we get

$$X([Q]+1),\ X([Q]+2),\ \ldots,\ X([R]) \tag{8.2.6}$$

positive integers, each of which is $\leqslant \tau$. Apart from possibly τ of these numbers, they can all be grouped according to the scheme of Lemma 2. Let one such group be

$$X(\alpha_s+1),\ X(\alpha_s+2),\ \ldots,\ X(\alpha_s+n_s).$$

Among its terms, for example, $X(x_s) = n_s$. Now,

$$-\frac{1}{\tau} < n_s f'(n_s) - Y(x_s) < \frac{1}{\tau}\ (0 < n_s \leqslant \tau),$$

so that

$$f'(x_s) = \frac{Y(x_s)}{n_s} + \frac{\theta}{\tau n_s}\ (-1 < \theta < 1).$$

We now consider the sum

$$\Omega_s = \sum_{x>\alpha_s+1}^{x\leqslant\alpha_s+n_s} \{f(x)\}.$$

Putting $x = x_s + z$, we find that

$$\Omega_s = \sum_{z=\alpha_s-x_s+1}^{z=\alpha_s-x_s+n_s} \left\{f(x_s) + zf'(x_s) + \frac{z^2}{2} f''[x_s + z\varrho(z)]\right\}$$
$$(-1 < \varrho(z) < 1).$$

But $-n_s < z < n_s$, because

$$z = x - x_s;\quad \alpha_s + 1 \leqslant x_s \leqslant \alpha_s + n_s,$$
$$\alpha_s + 1 \leqslant x \leqslant \alpha_s + n_s.$$

Therefore,

$$\Omega_s = \sum_{z=\alpha_s-x_s+1}^{z=\alpha_s-x_s+n_s} \left\{ f(x_s) + zf'(x_s) + \frac{z^2}{2} f''[x_s + z\varrho(z)] \right\}$$

$$= \sum_{z=\alpha_s-x_s+1}^{z=\alpha_s-x_s+n_s} \left\{ \frac{n_s f(x_s) + zY(x_s) + \frac{\theta}{\tau} z + \sigma(z) \frac{n_s^8}{2A}}{n_s} \right\},$$

since

$$f'(x_s) = \frac{Y(x_s)}{n_s} + \frac{\theta}{\tau n_s};$$

here, $|\theta| < 1$, $|\sigma(z)| < 1$ and $\sigma(z)$ retains the same sign. The expression

$$\frac{z}{2} f''(x_s + z\varrho(z))$$

lies between $z/2kA$ and $z^2/2A$, so that it is equal to

$$\frac{1}{k'A} \frac{n_s^2}{2} \left(\frac{z}{n_s}\right)^2 (k' \geqslant 1).$$

On denoting $(1/k')\,(z/n_s)^2 = \tau(z)$, we see that $\sigma(z) \geqslant 0$ and $\leqslant 1$, whence we obtain our formula. Here $z = \alpha_1 - x_s + 1, \ldots, \alpha_s - x_s + n_s$.

On putting

$$M = \alpha_s - x_s; \quad X = n_s; \quad Y = Y(x_s); \quad P(z) = n_s f(x_s) + \frac{\theta}{\tau} z + \sigma(z) \frac{n_s^3}{2A}; \quad C = \frac{n_s}{\tau} + \frac{n_s^3}{2A},$$

we arrive at the conditions of Lemma 3. On applying this lemma, we obtain

$$\left| Q_s - \frac{n_s}{2} \right| < \frac{n_s}{\tau} + \frac{n_s^3}{2A} + \frac{1}{2}.$$

We write down these inequalities for all the groups, while we write the inequality $-\frac{1}{2} \leqslant \{f(x)\} - \frac{1}{2} < \frac{1}{2}$ for all the x of the series

$$x = [Q] + 1, \quad [Q] + 2, \ldots, [R], \tag{8.2.7}$$

which appear in none of the groups. There are $< \tau$ of these x. We add these inequalities term by term, replacing $[R] - [Q]$ by $R - Q + \theta$; $|\theta| < 1$. We now have

$$\left| \sum_{x>Q}^{x \leqslant R} \{f(x)\} = \frac{1}{2} (R - Q) \right| < H,$$

where

$$H = \sum_s \left(\frac{n_s}{\tau} + \frac{n_s^3}{2A} + \frac{1}{2} \right) + \frac{1}{2}\tau + \frac{1}{2};$$

$\sum_s$ denotes summation over all the groups. This sum splits into two:

$$\sum_s \left(\frac{n_s}{\tau} + \frac{n_s^3}{2A} \right) \quad \text{and} \quad \sum_s \frac{1}{2}.$$

The first sum is less than (since $n_s < \tau$)

$$\left(\frac{1}{\tau} + \frac{\tau^2}{2A} \right) \sum_s n_s \leqslant \left(\frac{1}{\tau} + \frac{\tau^2}{2A} \right) (R - Q + 1).$$

The second sum $\sum_s \frac{1}{2} = \frac{1}{2}T$, where T is the number of groups. Since $\tau > 4$ and $\tau < \sqrt{A}$, we have

$$H < \left(\frac{1}{\tau} + \frac{\tau^2}{2A} \right) (R - Q) + \tau + \frac{T}{2}. \tag{8.2.8}$$

(3) We make use of the fact that

$$T < \sum \frac{1}{X}, \tag{8.2.9}$$

where X runs over the values

$$X([Q] + 1), \quad X([Q] + 2), \ldots, X([R]). \tag{8.2.6}$$

Let us consider how many times the same value X is repeated in the series (8.2.6).

With each given X there can be connected only those values of Y which satisfy the inequalities

$$-\frac{1}{\tau} < n_s f'(x_s) - Y(x_s) < \frac{1}{\tau}, \quad -\frac{1}{\tau} < Xf'(x) - Y < \frac{1}{\tau}.$$

Hence

$$Y > -\frac{1}{\tau} + n_s f'(x_s);$$

$Y > -1/\tau + XK$, K is the least number of the series (8.2.5),

$$Y < Xf'(x) + \frac{1}{\tau} < X\left(K + \frac{R - Q}{A} \right) + \frac{1}{\tau},$$

$$XK - \frac{1}{\tau} < Y < X\left(K + \frac{R - Q}{A} \right) + \frac{1}{\tau}.$$

The number of these values Y does not exceed

$$X\frac{R-Q}{A}+\frac{3}{2}.$$

This tells us how many Y there can be for one X. Only those numbers of the series (8.2.5) which satisfy

$$\frac{Y-\frac{1}{\tau}}{X}<f'(x)<\frac{Y+\frac{1}{\tau}}{X}$$

can be connected with any given pair of X and Y. In view of what was said in (1) ($f''(x) > 1/kA$), the number of such values is not greater than

$$\frac{2kA}{\tau X}+1<3k\frac{A}{\tau X}, \quad \text{since} \quad \tau<\sqrt{A},\ X<\tau.$$

Thus, not more than

$$\left(X\frac{R-Q}{A}+\frac{3}{2}\right)3k\frac{A}{\tau X}$$

numbers of the series (8.2.5): $f'([Q]+1), \ldots, f'([R])$, are connected with the given X. Hence the given X is encountered not more than this number of times in the series (8.2.6):

$$X([Q]+1), \ldots, X([R]).$$

For, $X(L)$ in this series is the value of X, chosen for $f'(L)$ so that

$$-\frac{1}{\tau}X(L)f'(L)-Y<\frac{1}{\tau}\ (0<X<\tau),$$

i.e. X is connected with the value $f'(L)$, while we have calculated how many of these values are connected with X. Consequently, there are not more than

$$k\left(3\frac{R-Q}{\tau}+\frac{9}{2}\frac{A}{\tau X}\right)$$

of these values. Hence, from (8.2.9),

$$\begin{aligned}T&<k\sum_{X>0}^{X<\tau}\left(3\frac{R-Q}{\tau}\frac{1}{X}+\frac{9}{2}\frac{A}{\tau}\frac{1}{X^2}\right)\\&<k\left(\frac{3}{2}\frac{R-Q}{\tau}\ln A+3\frac{R-Q}{\tau}+\frac{15}{2}\frac{A}{\tau}\right),\end{aligned}$$

since $\Sigma 1/n^2 = \pi^2/6 < 5/3$. We obtain from (8.2.8):

$$H < \left(\frac{1}{\tau} + \frac{\tau^2}{2A}\right)(R-Q) + \tau + \frac{1}{2}k\left(\frac{3}{2}\frac{R-Q}{\tau}\ln A + 3\frac{R-Q}{\tau} + \frac{15}{2}\frac{A}{\tau}\right) < (R-Q)\left(\frac{5}{2\tau} + \frac{\tau^2}{2A} + \frac{3}{4}k\frac{\ln A}{\tau}\right) + \frac{15}{4}k\frac{A}{\tau} + \tau.$$

Let $\tau = \sqrt{A \ln A}$; now $4 < \tau < \sqrt{A}$, and we have

$$\begin{aligned} H &< k\frac{(R-Q)}{A}\left(\frac{3}{4}\frac{A\ln A}{\tau}\right) + k\frac{2A}{\tau} \\ &\quad + (R-Q)\left(\frac{5}{2\tau} + \frac{\tau^2}{2A}\right) + \tau \\ &< 2k\left(\frac{R-Q}{A} + 1\right)(A\ln A)^{2/3} \qquad (A > 25). \end{aligned}$$

This proves the fundamental formula.

3. Application of Sonin's Formula

We have obtained the fundamental formula for the sum of the fractional parts

$$\sum_{x>Q}^{x\leqslant R} \{f(x)\} = \frac{R-Q}{2} + G, \tag{8.3.1}$$

where

$$|G| < 2k\left(\frac{R-Q}{A} + 1\right)(A\ln A)^{2/3}. \tag{8.3.2}$$

Hence

$$\sum_{x>Q}^{x\leqslant R} [f(x)] = \sum_{x>Q}^{x\leqslant R} f(x) - \frac{R-Q}{2} - G. \tag{8.3.3}$$

To compute the first term on the right-hand side, we use Sonin's formula (see [23]).

Let

$$\varrho(x) = [x] - x + \tfrac{1}{2}; \quad \sigma(x) = \int_0^x \varrho(x)\,dx.$$

Now,

$$\sum_{x>a}^{x\leqslant b} f(x) = \int_a^b f(x)\,dx + \varrho(b)f(b) - \varrho(a)f(a) - \sigma(b)f'(b) + \sigma(a)f'(a) + \int_a^b f''(x)\,\sigma(x)\,dx. \tag{8.3.4}$$

At the same time,

$$|\varrho(x)| \leqslant \tfrac{1}{2}; \quad |\sigma(x)| \leqslant \tfrac{1}{8}.$$

If $f''(x)$ does not change sign, then in view of the fact that $\sigma(x)$ does not change sign, and assuming that $|f'(x)| \leqslant \mu$ in the interval $[Q, R]$, we have

$$\left| \int_a^b f''(x)\, \sigma(x)\, dx \right| < \frac{\mu}{4}. \tag{8.3.5}$$

In addition, of course,

$$|\sigma(b)\, f'(b)| \leqslant \frac{\mu}{8} \quad \text{and} \quad |\sigma(a)\, f'(a)| \leqslant \frac{\mu}{8}.$$

Hence (8.3.4) and (8.3.3) yield

$$\sum_{x>Q}^{x\leqslant R} [f(x)] = \int_Q^R f(x)\, dx + \varrho(R)\, f(R) - \varrho(Q)\, f(Q) - \frac{R-Q}{2} + H, \tag{8.3.6}$$

where

$$|H| < \frac{\mu}{2} + 2k\left(\frac{R-Q}{A} + 1\right)(A \ln A)^{2/3}. \tag{8.3.7}$$

As an example of the application of Theorem 8.2.1 we consider the family of parabolas depending on the parameters n, a:

$$y = \frac{n + x^2}{a}, \quad n > 0; \quad -a \leqslant Q < R \leqslant a.$$

Here we have:

$$y' = \frac{2x}{a}, \quad \mu = 2, \quad y'' = \frac{2}{a}, \quad k = 1, \quad A = \frac{a}{2}.$$

We obtain for the number T of integral points inside the contour:

$$T = \frac{1}{3a}(R^3 - Q^3) + \varrho(R)\frac{n + R^2}{a} - \varrho(Q)\frac{n + Q^2}{a} + \frac{n}{a}(R - Q) - \frac{R-Q}{2} + H;$$

$$|H| \leqslant 1 + 10\left(\frac{a}{2}\ln\frac{a}{2}\right)^{2/3} < 8(a \ln a)^{2/3} \qquad \left(\frac{a}{2} > 25\right).$$

4. Generalization of Theorem 8.2.1

Let $r_{\max}/r_{\min} < \sigma$ on the arc of a curve, where r is the radius of curvature. Now $f''(x)$ does not change sign; $f'(x)$ is monotonic. Let it be increasing.

(1) Let $|f'(x)| \leqslant 1$ throughout the arc; $f'(x) \neq 0$. We have

$$y'' = \frac{(1+y'^2)^{3/2}}{r};$$

hence

$$\frac{1}{r} < f''(x) < \frac{2^{3/2}\sigma}{r}, \quad r = r_{\max}.$$

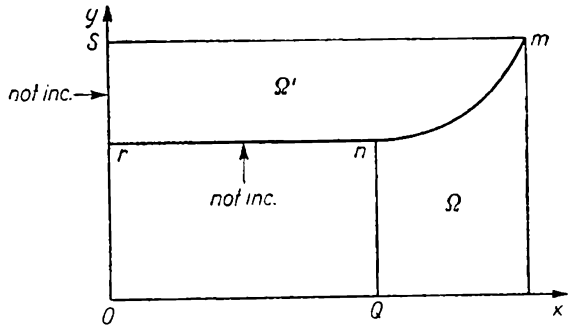

FIG. 6.

In this case, in the notation of § 3: putting $\mu = 1$; $A = r/2^{3/2}\sigma$; $k = 2^{3/2}\sigma$. Further, the fact that $y_1' - y_0' = \int_Q^R y''\,dx$ gives us

$$(R-Q)\frac{1}{r} < 2; \quad R - Q < 2r.$$

Now, if T is the number of integral points; S is the area enclosed by the curve, we have

$$T = S + \varrho(R)\,f(R) - \varrho(Q)\,f(Q) - \frac{R-Q}{2} + H \qquad (8.4.1)$$

and when $r/2^{3/2}\sigma > 25$,

$$|H| < \frac{1}{2} + 2^{5/2}\sigma(2^{5/2}\sigma + 1)\left(\frac{r}{2^{3/2}\sigma}\ln\frac{r}{2^{3/2}\sigma}\right)^{2/3} < 19\sigma^{4/3}(r\ln r)^{2/3}. \qquad (8.4.2)$$

(2) Let $|f'(x)| \geqslant 1$ everwhere on the arc (Fig. 6). $Or = k$; $OS = l$; on changing the role of the axes, we obtain for the region Ω' the number of integral points

$$T' = S' + \varrho(l)\, R - \varrho(h)\, Q - \tfrac{1}{2}(l - h) + H'';$$
$$|H''| < 19\sigma^{4/3}(r \ln r)^{2/3}.$$

The number of integral points in the rectangle $ORnS$ less the sides OR and OS is: $T + T' + [Q]\,[f(Q)] = [R]\,[f(R)]$, so that

$$T = [R]\,[f(R)] - [Q]\,[f(Q)] - T',$$
$$T' = S' + \varrho(f(R))\, R - \varrho(f(Q))\, Q + \tfrac{1}{2}(f(R) - f(Q)) + H'',$$
$$S' = Rf(R) - Qf(Q) - S,$$
$$-T' = S - Rf(R) + Qf(Q) - \varrho(f(R))\, R + \varrho(f(Q))\, Q$$
$$+ \tfrac{1}{2}(f(R) - f(Q)) - H''.$$

Hence we find that

$$\begin{aligned} T &= [R]\,[f(R)] - [Q]\,[f(Q)] + S - Rf(E) + Qf(Q) \\ &\quad - \varrho(f(R))\, R + \varrho(f(Q))\, Q + \tfrac{1}{2}(f(R) - f(Q)) - H'' \\ &= S + f(R)\,([R] - R + \tfrac{1}{2}) - f(Q)\,([Q] - Q + \tfrac{1}{2}) \\ &\quad - \frac{R - Q}{2} - [R]\, f(R) - R([f(R)] - f(R)) + [R]\,[f(R)] \\ &\quad + [Q]\, f(Q) + Q([f(Q)] - f(Q)) - [Q]\,[f(Q)] - H'' \\ &= S + \varrho(R)\, f(R) - \varrho(Q)\, f(Q) - (R - [R])\,(f(R) - [f(R)]) \\ &\quad + (Q - [Q])\,(f(Q) - [f(Q)]) - H'' \\ &= S + \varrho(R)\, f(R) - \varrho(Q)\, f(Q) + H''' ; \\ &\qquad |H'''| < 1 + 19\sigma^{4/3}(r \ln r)^{2/3} < 20\sigma^{4/3}(r \ln r)^{2/3}. \end{aligned} \tag{8.4.3}$$

(3) Let $|f'(x)| \leqslant 1$ at some points of the arc nm, and > 1 at other points. The arc nm can be divided into not more than 3 parts, in each of which $|f'(x)| \leqslant 1$ or $\geqslant 1$. Since $f'(x)$ is monotonic on it, say increasing, if it is initially > 0, we obtain a division into 1 or 2 arcs. If it is initially < 0 and < -1, we obtain a division into 3, 2 or 1 arcs, if it is < 0 and > -1, into 2 or 1 arcs. The region Ω splits into $\leqslant 3$ disjunctive regions. We apply (8.4.2) or (8.4.3) to each region and add:

$$T = S + \varrho(R) - \varrho(Q)\, f(Q) - \tfrac{1}{2}(R - Q) + \Gamma; \tag{8.4.4}$$
$$|\Gamma| < 60\sigma^{4/3}(r \ln r)^{2/3}.$$

5. Distribution over a Closed Contour

Let Ω, of area S, be bounded by the contour C (Fig. 7); $r_{\max}/r_{\min} < \sigma$. We draw the two vertical tangents. Ω' is the shaded, and Ω'' the lower region; γ is the number of points with integral coordinates on the arc *nrm*. By (8.4.4): if T is the number of points inside the contour, we have

$$T + \gamma = S + B(r \ln r)^{2/3}.$$

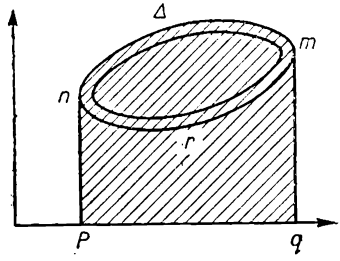

FIG. 7.

To eliminate γ, we take a homothetic but very close contour, for which there are no integral points on the upper arc. Hence

$$T = S + B(r \ln r)^{2/3}, \quad r = r_{\max}.$$

As an example, we consider the ellipse:

$$ax^2 + 2bxy + cy^2 \leqslant M,$$
$$b^2 - ac = -D < 0,\ a > 0.$$

If the initial ellipse $ax^2 + 2bxy + cy^2 \leqslant 1$, the given ellipse is obtained by a $\sqrt{M}$ times extension; its curvature is $\sqrt{M}$ times less, $\sigma = r_{\max}/r_{\min} = \text{const}$, $r = \varrho\sqrt{M}$; the area of Ω is $S = \pi M/\sqrt{D}$. We have from the above:

$$T' = \frac{\pi M}{\sqrt{D}} + BM^{1/3}(\ln M)^{2/3}.$$

Notice that Vinogradov's method, as described here, in spite of its elementary nature, yields error estimates which can in the general case be improved by not more than the factor $(\ln r)^{2/3}$. This improvement can be achieved by using complicated analytic methods. Landau[41] used analytic methods to prove the following

theorem: let $w > 0$ be an integer; for $\frac{1}{2} \leqslant u \leqslant w$, let a family of functions $f(u)$ be given, depending on the parameter $A \to \infty$ and such that $f''(u)$ exists and is continuous; $f(\frac{1}{2}) > 2$; $0 < f'(u) < 1$; $f''(u) > 1/A (A > 1)$.

Now, if S is the area under the segment of curve in the interval $[1, w]$, and T is the number of integral points in this area, excluding the axes, we have

$$|T - S| < cA^{2/3}. \tag{8.5.1}$$

This is a $(\ln r)^{2/3}$ times improvement on Vinogradov's result discussed above. Further improvements are impossible in the general case. This follows from Jarnik's result of 1926: there exists a curve L, depending on a parameter A and satisfying the above-mentioned conditions, for which

$$|T - S| > c_1 A^{2/3}. \tag{8.5.2}$$

For special types of contour, such as the circle or ellipse, better results than (8.5.1) may be obtained, although, as remarked in § 1, the strict order of $T - S$ has not yet been found.

The condition of § 4 for the ratio of the radii of curvature is not fulfilled for a hyperbola, which corresponds to the divisor problem. However, a modification of Vinogradov's method is possible here, consisting in dividing the contour into pieces, inside each of which the relative variation of the radius of curvature is permissible (see Vinogradov[1]).

In this way, for instance, the ratio

$$\varrho(n) = Bn^{1/3}(\ln n)^{5/3}$$

is obtained in (8.1.9) for the divisor problem.

The elementary method that we have described for calculating the integral points in a contour is also applicable to three-dimensional problems, for instance to the problem of computing the points inside and on the sphere

$$x^2 + y^2 + z^2 = R^2.$$

CHAPTER 9

THE DISTRIBUTION OF POWER RESIDUES

1. A Theorem of Vinogradov

In this chapter we shall consider the distribution of power residues and nonresidues of order $l > 1$ for a large prime modulus $p \equiv 1 \pmod{l}$. We shall start by deseribing Vinogradov's elementary method (see [4]).

THEOREM 9.1.1. (Vinogradov.) *Let $p > 2$ be prime, $l > 1$, $p \equiv 1 \pmod{l}$ and the numbers $1, 2, \ldots, p-1$ be divided into residue and nonresidue classes $A_0, A_1, \ldots, A_{l-1}$ of order l modulo p.*

Let a be not divisible by p; b be an arbitrary integer and $0 < h < p$.

The number c_i of numbers of the ith class A_i $(i = 0, 1, \ldots, l-1)$, contained in the progression $ax + b$ $(x = 0, 1, \ldots, h-1)$ is

$$c_i = \frac{h}{l} + \Delta_i, \tag{9.1.1}$$

where

$$\Delta_i^2 < 2T.$$

Here,

$$T = \sum_{x=1}^{h} \sum_{(y)} \left(\frac{p}{xy} + 1 \right) \tag{9.1.2}$$

and the summation over y for a given x is over all $y \leqslant p/x$; $(y, x) = 1$.

We observe that (9.1.2) implies

$$T = Bp(\ln p)^2, \tag{9.1.3}$$

so that

$$\Delta_i = B\sqrt{p}\ln p. \tag{9.1.4}$$

The proof of Theorem 9.1.1 is based on the following three lemmas.

LEMMA 1. *Let A, B, be real numbers, $k > 0$ be an integer;*

$$A = \frac{t}{q} + \frac{\theta}{qk} \quad (0 < q \leqslant k,\ (t, q) = 1,\ |\theta| < 1);$$

c, m are integers; $0 < m \leqslant k/q$. *Now,*

$$\sum_{x=c}^{c+mq-1} \{Ax + B\} = \tfrac{1}{2} mq + \tfrac{1}{2} \varrho(m + 1); \quad |\varrho| \leqslant 1. \qquad (9.1.5)$$

The proof of this lemma is very similar to the proof of the lemma of Chapter 7.

We replace A by its above value and find that $Ax + B = (tx + f(x))/q$, where $f(x) = Bq + \theta x/k$; $f(x)$ is a linear function of x, so that, when $c \leqslant x \leqslant c + mq - 1$, $f(x)$ lies between the extreme values $f(c)$ and $f(c + mq - 1)$. The difference between these values has an absolute value not greater than $|(mq - 1)/k| \leqslant 1$. Thus the integers $[f(c)]$ and $[f(c + mq - 1)]$ differ by not more than unity.

We divide the sum on the left-hand side of (9.1.5) into m sums:

$$S = S_0 + S_1 + \cdots + S_{m-1}; \quad . S_i = \sum_{c+iq}^{c+iq+q-1} \{Ax + B\}.$$

Let us consider one such sum S_j. If $[f(c + jq)] = [f(c + jq + q - 1)] = n$, on putting $f(x) = n + \lambda(x)$ $(0 \leqslant \lambda(x) < 1$ when $c + jq \leqslant x \leqslant c + jq + q - 1)$, we have

$$S_j = \sum_{c+jq}^{c+jq+q-1} \left\{\frac{tx + n + \lambda(x)}{q}\right\} = \sum_{r=0}^{q-1} \frac{r + \lambda(x)}{q}, \qquad (9.1.6)$$

since $tx + n$ runs along with x over the complete system of residues (mod q). We obtain from (9.1.6):

$$S_j = \frac{q}{2} + \frac{\varrho_j}{2} \qquad (|\varrho_j| \leqslant 1). \qquad (9.1.7)$$

If the numbers $[f(c + jq)]$ and $[f(c + jq + q - 1]$ differ by unity, we denote the lesser of them by n and put $f(x) = n + \lambda(x)$; we now have for the values of x in question: $0 \leqslant \lambda(x) < 2$, where the extreme values of $\lambda(x)$ must lie in the intervals $[0, 1)$ and $[1, 2)$. We find in the same way as above that

$$S_j = \sum_{r=0}^{q-1} \left\{\frac{r + \lambda(x)}{q}\right\}$$

where

$$0 \leqslant r \leqslant q - 2, \left\{\frac{r + \lambda(x)}{q}\right\} = \frac{r + \lambda(x)}{q}.$$

On recalling that $r = q - 1$ and putting $\delta = 0$ or $\delta = 1$, we find that

$$S_j = \sum_{r=0}^{q-1} \frac{r + \lambda(x)}{q} - \delta = \frac{q-1}{2} - \delta + \frac{1}{q} \sum_{c+jq}^{c+jq+q-1} \lambda(x)$$

$$= \frac{q}{2} - \frac{1}{2} - \varepsilon + \frac{1}{2}(\lambda(c + jq) + \lambda(c + jq + q - 1)), \quad (9.1.8)$$

whence

$$S_j = \frac{q}{2} + \varrho_i \qquad (|\varrho_i| \leqslant 1). \quad (9.1.9)$$

We now observe that there can only be one value of j for which $[f(c + jq)] = [f(c + jq + q - 1)]$. On adding the expressions obtained for S_j, we arrive at (9.1.5).

LEMMA 2. *Let $p > 2$ be prime; α one of the numbers $1, 2, \ldots, p - 1$; β_α an integer which can vary as α varies, $h \in [1, p - 1]$ an integer.*

We put

$$S_\alpha = \sum_{x=0}^{h-1} \left\{\frac{\alpha x + \beta_\alpha}{p}\right\}, \quad L_\alpha = S_\alpha - \frac{h}{2}. \quad (9.1.10)$$

Now,

$$\sum_{\alpha=1}^{p-1} |L_\alpha| < T,$$

where T is given by (9.1.2).

To investigate the sum S_α, we find the irreducible fraction t_0/x_0 under the conditions

$$\frac{\alpha}{p} = \frac{t_0}{x_0} + \frac{\theta_0}{x_0 h}; \quad x_0 \in (0, h]; |\theta_0| < 1.$$

We put

$$m_0 = \left[\frac{h}{x_0}\right], \quad h_1 = h - m_0 x_0 (m_0 > 0, \quad 0 < h_1 < x_0),$$

and divide S_α into two parts:

$$S_\alpha = \sum_{x=0}^{m_0 x_0 - 1} \left\{\frac{\alpha x + \beta_\alpha}{p}\right\} + \sum_{m_0 x_0}^{h-1} \left\{\frac{\alpha x + \beta_\alpha}{p}\right\}.$$

We apply Lemma 1 to the first part, putting $q = x_0$. Using (9.1.5) where we take $m_0 x_0 = h - h_1$, $m_0 \leqslant h/x_0$, we obtain the expression

$$S_\alpha = \frac{1}{2}(h - h_1) + \frac{1}{2}\varrho_0\left(\frac{h}{x_0} + 1\right) + S'_\alpha, \quad |\varrho_0| \leqslant 1,$$

where

$$S'_\alpha = \sum_{x=c_0}^{c_0+h_1-1} \left\{\frac{ax+\beta_a}{p}\right\} \quad (c_0 = m_0 x_0)\,.$$

We proceed in the same way with the new sum S'_α. We put

$$\frac{\alpha}{p} = \frac{t_1}{x_1} + \frac{\theta_1}{x_1 h_1}\,; \quad 0 < x_1 \leqslant h_1; \quad (t_1, x_1) = 1, \quad |\theta_1| < 1\,;$$

$$m_1 = \left[\frac{h_1}{x_1}\right], \quad h_2 = h_1 - m_1 x_1, \quad 0 < h_2 < x_1\,,$$

and as before, obtain the expression

$$S'_\alpha = \frac{1}{2}(h_1 - h_2) + \frac{1}{2}\varrho_2\left(\frac{x_1}{h_1} + 1\right) + S''_\alpha, \quad |\varrho_2| \leqslant 1\,;$$

$$S''_\alpha = \sum_{x=c_1}^{c_1+h_2-1} \left\{\frac{ax+\beta_\alpha}{p}\right\}, \qquad c_1 = m_0 x_0 + m_1 x_1\,.$$

We continue this process; it will be seen that h, h_1, h_2, ... are decreasing, so that at some stage $h_{n+1} = 0$, and the process terminates. The last equation has the form

$$S^{(n)}_\alpha = \frac{1}{2} h_n + \frac{1}{2}\varrho_n\left(\frac{h_n}{x_n} + 1\right), \quad |\varrho_n| \leqslant 1\,,$$

where

$$\frac{\alpha}{p} = \frac{t_n}{x_n} + \frac{\theta_n}{x_n h_n}\,, \quad 0 < x_n \leqslant h_n, \quad (t_n, x_n) = 1, \quad |\theta_n| < 1\,,$$

$$m_n = \left[\frac{h_n}{x_n}\right] = \frac{h_n}{x_n}\,; \quad h_{nn} = h_n - m_n x_n = 0\,.$$

On adding the expressions obtained for S_α, S'_α, ..., $S^{(n)}_\alpha$, we find that

$$S_\alpha = \frac{1}{2} h + \frac{1}{2}\varrho_0\left(\frac{h}{x_0} + 1\right) + \frac{1}{2}\varrho_1\left(\frac{h_1}{x_1} + 1\right)$$

$$+ \cdots + \frac{1}{2}\varrho_n\left(\frac{h_n}{x_n} + 1\right),$$

whence

$$L_\alpha = S_\alpha - \frac{h}{2} = \frac{\varrho}{2}\left[\left(\frac{h}{x_0} + 1\right) + \left(\frac{h_1}{x_1} + 1\right)\right.$$

$$\left. + \cdots + \left(\frac{h_n}{x_n} + 1\right)\right] (|\varrho| \leqslant 1)$$

and

$$\sum_{\alpha=1}^{p-1} |L_\alpha| \leqslant \frac{1}{2} \sum_{\alpha=1}^{p-1} \left[\left(\frac{h_0}{x^0} + 1 \right) + \left(\frac{h_1}{x_0} + 1 \right) + \cdots + \left(\frac{h_n}{x_n} + 1 \right) \right]. \tag{9.1.11}$$

In the last sum the numbers x_i, h_i depend on α and are determined by it.

We now collect all the terms in the sum of the form $(h_i/x_i) + 1$, where x has the same value. We write the equation

$$\frac{\alpha}{p} = \frac{t_i}{x_i} + \frac{\theta_i}{x_i h_i}$$

in the form

$$\alpha x_i = p t_i + y_i, \qquad y_i = \frac{p\theta_i}{h_i}.$$

We see from this that y_i is an integer; $(y_i, x_i) = 1$ and $\alpha x_i \equiv y_i \pmod{p}$ $(0 < |y_i| < p/h_i)$. If x_i, y_i are given and satisfy the given conditions, then t_i is defined (for the fixed α, h_i).

To obtain an upper bound for the terms $(h_i/x_i) + 1$, where the x_i are equal to the given number x $(0 < x \leqslant h)$, we consider the corresponding values of α, y_i. Since $|y_i| \in (0, p_i/h_i)$, y_i can take only the values $\pm 1, \pm 2, \ldots, \pm[p/x]$ relatively prime to x. Once y_i is chosen, α is determined from the congruence $\alpha x_i \equiv y_i \pmod{p}$. Further, $h_i/x_i < p/x_i|y_i|$, so that the sum of these terms $(h_i/x_i) + 1$ does not exceed $2\Sigma[(p/xy) + 1]$, where the summation is over all $y = 1, 2, \ldots, [p/x]$, relatively prime to x. This proves Lemma 2.

LEMMA 3. *Let α, p, β_α be defined as in the previous lemma; h, γ are integers $\in (0, p)$. Let R_α be the number of numbers $ax + \beta_\alpha$ $(x = 0, 1, 2, \ldots, h-1)$, for which*

$$0 < \left\{ \frac{ax + \beta_a}{p} \right\} < \frac{\gamma}{p}.$$

We put

$$H(\alpha, \beta_\alpha, h, \gamma) = R_\alpha - \frac{h\gamma}{p}. \tag{9.1.12}$$

Now,

$$\sum_{\alpha=1}^{p-1} |H(\alpha, \beta_\alpha, h, \gamma)| < 2T. \tag{9.1.13}$$

This lemma is easily deduced from Lemma 2. We observe that

$$\left\{ \frac{\alpha x + \beta_\alpha - \gamma}{p} \right\} - \left\{ \frac{\alpha x + \beta_\alpha}{p} \right\} = 1 - \frac{\gamma}{p},$$

if the least residue of $\alpha x + \beta_\alpha$ modulo p lies in the interval $(0, \gamma]$ and is γ/p if it lies in the interval $[\gamma, p)$. In view of this, we have

$$S'_\alpha - S_\alpha = \sum_{x=0}^{h-1} \left\{\frac{\alpha x + \beta_\alpha - \gamma}{p}\right\} - \sum_{x=0}^{h-1} \left\{\frac{\alpha x + \beta_\alpha}{p}\right\} = H(\alpha, \beta_\alpha, h, \gamma).$$

We have by Lemma 2:

$$\sum_{\alpha=1}^{p-1} \left| S'_\alpha - \frac{h}{2} \right| < T; \quad \sum_{\alpha=1}^{p-1} \left| S_\alpha - \frac{h}{2} \right| < T,$$

whence

$$\sum_{\alpha=1}^{p-1} |S'_\alpha - S_\alpha| = \sum_{\alpha=1}^{p-1} |H(\alpha, \beta_\alpha, h, \gamma)| < 2T,$$

which proves Lemma 3.

2. Proof of Theorem 9.1.1

We turn to the proof of Theorem 9.1.1. Let j be one of the numbers $0, 1, 2, \ldots, l-1$. We form the $(p-1)\,h : l$ products

$$\alpha(ax + b), \tag{9.2.1}$$

where α runs over all the numbers of the class A_i, and $x = 0, 1, 2, \ldots, h-1$ independently of α. For each of these products we define a number u by the congruence:

$$au + b \equiv a(ax + b) \pmod{p} \quad (0 \leqslant u < p). \tag{9.2.2}$$

We denote by D the number of products (9.2.1) for which the corresponding value of u is less than h. This number D will be computed by two different methods; the resulting equality will yield the proof that we require. We first consider the products (9.2.1) which correspond to the same value of α. Let $aa' \equiv 1 \pmod{p}$, in which case the congruence (9.2.2) can be replaced by $u \equiv \alpha x + \beta_\alpha$, $\beta_\alpha = \alpha b a' - b a'$ and is independent of x. On applying Lemma 3 for the given value of α and summing over all α, we obtain

$$D = \frac{p-1}{l}\frac{h^2}{p} + \sum_\alpha H(\alpha, \beta_\alpha, h, h). \tag{9.2.3}$$

We now compute D by another method, while leaving x constant in the products (9.2.1). If now $ax + b \equiv 0 \pmod{p}$, then all $(p-1)/l$ of the corresponding products must belong to D. If

$ax + b \equiv 0 \pmod{p}$, this number is one of the c_j numbers of class A_j contained in our progression. In this case, by (9.2.2), $au + b$ belongs to the class A_{i+j} (the number $i + j$ must be taken modulo l). In view of this, the numbers α include c_{i+j} such that $u \in (0, h)$. Hence

$$D = c_0c_i + c_1c_{i+1} + \cdots + c_{l-1}c_{i+l-1} + \eta, \tag{9.2.4}$$

where $\eta = (p-1)/l$, if the congruence $ax + b \equiv 0 \pmod{p}$ $(0 \leqslant x \leqslant h-1)$ is solvable, and $\eta = 0$ otherwise.

On comparing (9.2.3) and (9.2.4), we obtain for any i:

$$c_0c_i + c_1c_{i+1} + \cdots + c_{l-1}c_{i+l-1} + \eta = \frac{p-1}{l}\frac{h^2}{p} + \sum_{\alpha} H(\alpha, \beta_\alpha, h, h), \tag{9.2.5}$$

where the summation on the right is over $\alpha \in A_i$. We put $i = 0$ in (9.2.5), and subtract (9.2.5) term by term from the resulting expression. Since the numbers c_i, c_{i+1}, ..., c_{i+l-1} differ only in their order from c_0, c_1, ..., c_{l-1}, we obtain

$$(c_0 - c_i)^2 + (c_1 - c_{i+1})^2 + \cdots + (c_{l-1} - c_{i+l-1})^2 = 2\sum_{\alpha'} H(\alpha', \beta_{\alpha'}, h, h) - 2\sum_{\alpha} H(\alpha, \beta_\alpha, h, h), \tag{9.2.6}$$

where α, α' on the right-hand side run over all the number of the classes A_i and A_0. By Lemma 3, the right-hand side does not exceed $4T$; hence, for any r,

$$(c_r - c_{r+i})^2 + (c_r - c_{r-i})^2 < 4T. \tag{9.2.7}$$

Thus

$$2\sum_{i=0}^{l-1} (c_r - c_i)^2 < 4T(l-1). \tag{9.2.8}$$

We now put

$$\frac{c_0 + c_1 + \cdots + c_{l-1}}{l} = \bar{c}, \quad c_i = \bar{c} + \delta_i,$$

in which case we have

$$\sum_{i=0}^{l-1} (\delta_r - \delta_i)^2 < 2T(l-1);$$

since $\sum_{i=0}^{i-1} \delta_i = 0$, we have

$$l\delta_r^2 + \sum_{i=0}^{l-1} \delta_i^2 < 2T(l-1),$$

whence

$$\delta_r^2 < 2T\frac{l-1}{l+1}. \tag{9.2.9}$$

If the given progression $ax + b$ $(x = 0, 1, \ldots, h-1)$ does not contain zero, then obviously,

$$\bar{c} = \frac{h}{l}; \quad \delta_r = \Delta_r, \quad \Delta_r^2 < 2T\frac{l-1}{l+1} < 2T.$$

If it does contain zero modulo ϱ, then

$$\bar{c} = \frac{h-1}{l}; \quad \Delta_r = \delta_r - \frac{1}{l}; \quad \Delta_r^2 < 2T\frac{l-1}{l+1} + \frac{2|\delta_r|}{l};$$

further, $\delta_i = c_i - \bar{c}$, so that $|\delta_i| < p/2 < T/2$ and

$$\Delta_r^2 < 2T\left(\frac{l-1}{l+1} + \frac{1}{2l}\right) < 2T,$$

which completes the proof of the theorem.

3. Other Elementary Theorems on the Distribution of Characters. Unsolved Problems

Theorem 9.1.1 has a simple statement in the language of the arithmetical theory of characters (see, for example [3]). Let $p \equiv 1 \pmod{l}$ be a prime, and $\chi(m)$ a non-principal character of the lth power modulo p. We now have

$$\left|\sum_{x=0}^{h-1} \chi(ax+b)\right| < 2lT. \tag{9.3.1}$$

A less precise statement (using (9.1.4)) is

$$\left|\sum_{x=0}^{h-1} \chi(ax+b)\right| = Blp^{1/2}\ln p. \tag{9.3.2}$$

Notice that the case of the distribution of characters in a progression $ax + b$ easily reduces to the case of their distribution in the series of consecutive numbers $x = 1, 2, \ldots, H$ $(H \leqslant p-1)$. For, since a is chosen $\not\equiv 0 \pmod{p}$, then $\chi(ax+b) = \chi(a)\chi(x+ba')$ where $aa' \equiv 1 \pmod{p}$. If we know that

$$\left|\sum_{y=1}^{Y} \chi(y)\right| = Blp^{1/2}\ln p, \tag{9.3.3}$$

then we also have

$$\left|\sum_{x=0}^{h-1} \chi(x + ba')\right| = Blp^{1/2} \ln p, \tag{9.3.4}$$

whence (9.3.2) follows.

We shall be concerned in future with the non-principal characters of the second power (quadratic characters), which have the simplest structure. Given any prime $p > 2$, these characters $\chi_p(m)$ are well known to be the same as the Legendre symbols (m/p). They are characterized by the following properties:

(1) $\chi_p(n)$ is defined on the integers $0, \pm 1, \pm 2, \ldots$ and takes the values $0, +1$ or -1 on them;
(2) $\chi_p(n)$ is periodic: $\chi_p(n + p) = \chi_p(n)$;
(3) $\chi_p(n)$ is multiplicative: $\chi_p(mn) = \chi_p(m)\,\chi_p(n)$;
(4) $\chi_p(n)$ takes the value -1 for at least one n.

These four properties define the function $\chi_p(n)$ as the value of the Legendre symbol (m/p). The structure of the function $\chi_p(n)$ is simple. The function is of great value in analytic and algebraic theory of numbers, and it has been studied by many authors during the last 160 years or so. The first results, concerning its analytic behaviour as $p \to \infty$, were obtained by Gauss in 1796. Since that time relatively little has been discovered about the analytic properties of the function. There are still various unsolved problems which present great difficulties. Let us take the series of numbers:

$$\chi_p(1),\ \chi_p(2), \ldots, \chi_p(p-1). \tag{9.3.5}$$

Obviously, $\chi_p(1) = 1$, while the numbers $+1$ or -1 follow further on in the series. If $\chi_p(m) = +1$, m is referred to as a quadratic residue (mod p); if $\chi_p(m) = -1$, m is a quadratic nonresidue (mod p). If $m = \mu^2 < p$ (m is a square), then obviously m is a quadratic residue. The distance between the adjacent squares $\mu^2 < p$ and $(\mu + 1)^2 < p$ does not exceed $2\sqrt{p} + 1$; consequently, in the number series (9.3.5) the distance between adjacent quadratic residues likewise does not exceed $2\sqrt{p} + 1$. The question of the distance between adjacent quadratic nonresidues is extremely difficult. More than forty years ago, Vinogradov put forward a number of hypotheses concerning the distribution of quadratic characters, which we reproduce below.

HYPOTHESIS I. *The distance between adjacent quadratic nonresidues in the number series* (9.3.5) *is* $B_\varepsilon p^\varepsilon$, *for any* $\varepsilon > 0$ *and* $p \to \infty$.

In other words, if $d(p)$ is the maximum distance between adjacent nonresidues in the series (9.3.5), then with any given $\varepsilon > 0$,

$$\lim_{p\to\infty} \frac{d(p)}{p^\varepsilon} = 0. \tag{9.3.6}$$

A particular case, which is of special interest for the analytic theory of numbers, is the following hypothesis concerning the behaviour of the last quadratic nonresidue $N(p)$ in series (9.3.5).

HYPOTHESIS I'.

$$\lim_{p\to\infty} \frac{N(p)}{p^\varepsilon} = 0 \tag{9.3.7}$$

for any given $\varepsilon > 0$.

The question of the least quadratic residue has a trivial solution: the number 4 is the quadratic residue; if 2 and 3 are quadratic nonresidues, then $6 = 2 \cdot 3$ is a quadratic residue, and so on. However, if we pose the problem of the least prime quadratic residue $P(p)$, then this again is extremely difficult.

HYPOTHESIS II.

$$\lim_{p\to\infty} \frac{P(p)}{p^\varepsilon} = 0 \tag{9.3.8}$$

for any given $\varepsilon > 0$.

At the present time we are still a long way from proving these hypotheses; I is especially difficult. One of the basic hypotheses of modern analytic theory of numbers is the extension of Riemann's hypothesis for Dirichlet L-series. If this hypothesis is correct, then (9.3.7) and (9.3.8) are also correct in an even stronger form: $N(p) = B(\ln p)^2$; $P(p) = B(\ln p)^2$; but no progress is made towards Hypothesis I.

Various theorems on the interconnections between the above hypotheses and their connections with other hypotheses may be found in an article by Linnik and Renyi.[17] In the present section we shall also give the results obtained for these hypotheses by elementary methods.

THEOREM 9.3.1.

$$N(p) < \frac{1}{2} + \sqrt{p + \frac{1}{4}}\,. \tag{9.3.9}$$

Let $N(p) = n$. It is worth mentioning that n is a prime: if $n = n_1 n_2$, then $\chi_p(n_1) = -1$, or $\chi_p(n_2) = -1$; if $n_i \neq n$ or $n_i \neq 1$ $(i = 1, 2)$, a contradiction is obtained. We find the number $\xi \in (0, n)$ such that $p + \xi \equiv 0 \pmod{n}$. Since $\xi < n$, then $\chi_p(\xi) = \chi_p(p + \xi) = +1$, so that $\chi_p((p + \xi)/n) = -1$. But now $(p + \xi)/n \geqslant n$; and all the more $(p/n) + 1 > n$. Hence $n < \frac{1}{2} + \sqrt{p + 1/4}$, which it was required to prove.

THEOREM 9.3.2.

$$d(p) < 2\sqrt{p}. \tag{9.3.10}$$

Let there be a space of length greater than $2\sqrt{p}$ somewhere in series (9.3.5) between the signs (-1).

Let u be the left-hand prime of the space corresponding to the number series

$$1, 2, \ldots, p - 1. \tag{9.3.11}$$

We now have: $\chi_p(u + \xi) = +1$ $(\xi = 0, 1, \ldots, H)$, where $H > 2\sqrt{p}$. We consider the numbers $N(p)(u + \xi) = nu + n\xi$, where $n = N(p)$ is the least quadratic nonresidue mod p. We have: $\chi(nu + n\xi) = -1$ $(\xi = 0, 1, \ldots, H)$; by theorem (9.3.1), $n < 2\sqrt{p}$ $(p > 3)$. The distance between adjacent numbers $nu + n\xi \pmod{p}$ will obviously be less than $2\sqrt{p}$; if $H > 2\sqrt{p}$, then the number series (9.3.11) must contain quadratic nonresidues spaced less than $2\sqrt{p}$ apart, which is a contradiction. Thus (9.3.10) is proved.

The further investigation of the asymptotic formula for $N(p)$ by elementary methods requires employment of the elements of the theory of prime distribution. In 1926 Vinogradov proved the following.

THEOREM 9.3.3 (Vinogradov).

$$N(p) = Bp^{1/2\sqrt{e}} (\ln p)^2, \tag{9.3.12}$$

where e is the base of natural logarithms.

For the proof, we consider the number series $1, 2, \ldots, p_1$, where $p_1 = \sqrt{p}(\ln p)^2$. In accordance with (9.3.2), we have

$$\sum_{m \leqslant p_1} \chi_p(m) = B\sqrt{p} \ln p = B \frac{p_1}{\ln p}. \tag{9.3.13}$$

If $N(p) \leqslant p_1^{1/2}$, then obviously (9.3.12) is true, because $p_1^{1/2} = Bp^{1/2\sqrt{e}} (\ln p)^2$. We can therefore assume that $N(p) > p_1^{1/2}$. Let $m \leqslant p_1$; $\chi_p(m) = -1$. In this case there can be only one quadratic non-

residue among the prime divisors of m; if there were more, the inequality $N(p) > p_1^{1/2}$ could not be fulfilled. Thus, if $\chi_p(m) = -1$, the number m has the form $m = p\xi$, where $\chi_p(q) = -1$; $q \geqslant N(p) > p_1^{1/2}$, and $\chi_p(\xi) = +1$. We consider the sum

$$S = \sum_{\substack{\chi_p(m)=-1 \\ m \leqslant p_1}} 1. \tag{9.3.14}$$

By virtue of (9.3.13), we have

$$S = \frac{p_1}{2} + B\frac{p_1}{\ln p}. \tag{9.3.15}$$

On the other hand, it is obvious from what was said above that

$$S \leqslant S_1 = \sum_{\substack{m=q\xi \leqslant p_1 \\ N(p) \leqslant q \leqslant p_1}} 1, \tag{9.3.16}$$

where the summation is over the primes q and the positive numbers ξ. We have

$$S_1 = \sum_{N(p) \leqslant q \leqslant p_1} \left[\frac{p_1}{q}\right]. \tag{9.3.17}$$

If we omit the square brackets in (9.3.17), the total error will not exceed $\sum_{q \leqslant p_1} 1$. In accordance with what was said in Chapter 3, this error has the upper bound $Bp_1/(\ln p_1)$. Hence

$$S_1 = \sum_{N(p) \leqslant q \leqslant p_1} \frac{p_1}{q} + B\frac{p_1}{\ln p_1}. \tag{9.3.18}$$

We have further:

$$\sum_{q \leqslant x} \frac{1}{q} = \ln\ln x + C + \frac{x}{\ln x}, \tag{9.3.19}$$

where C is a constant. Hence

$$S_1 = p_1 \ln\frac{\ln p_1}{\ln N(p)} + \frac{Bp_1}{\ln p}. \tag{9.3.20}$$

On comparing with (9.3.15), we find that

$$\ln\frac{\ln p_1}{\ln N(p)} \geqslant \frac{1}{2} + \frac{B}{\ln p_1}, \tag{9.3.21}$$

$$\frac{\ln p_1}{\ln N(p)} \geqslant e^{1/2} + \frac{B}{\ln p_1}, \tag{9.3.22}$$

$$\ln N(p) \leqslant \frac{\ln p_1}{\sqrt{e}} + B. \tag{9.3.23}$$

In view of the definition of p_1, (9.3.12) now follows. Let us notice a further point, which is important for what follows. If the proof of Theorem 9.3.3 is followed carefully, it will be seen that the essential fact is not that $p_1 = \sqrt{p}(\ln p)^2$, but solely that $\sum_{m \leqslant p_1} \chi_p(m) = Bp_1/(\ln p)$. If this formula holds for any value of $p_1 = p_1(p)$, then the above arguments show that (9.3.23) is also true.

By using the modern theory of primes, and in particular the theorem on the least prime in a progression, due to one of the present authors, we are able to say something about the possible large values of $N(p)$. Thus Salie[44] and Friedländer[24] proved that there exist infinitely many primes p for which $N(p) > C_0 \ln p$. On the other hand, the numbers for which $N(p) > p^\varepsilon$, where the given $\varepsilon > 0$, are sparse, if they exist. We shall describe primes p for which $N(p) > p^\varepsilon$ as ε-exclusive. Linnik[11] showed that, given a sufficiently large N, the number of such numbers in the interval $[N^\varepsilon, N]$ does not exceed a constant: $C(\varepsilon) = 320\pi(g + 2)^g g!$, where $g = [2/\varepsilon^2 + 1]$. Friedländer[24] improved on this bound.

Notice also a curious fact: to prove hypothesis I′: $N(p) = B_\varepsilon p^\varepsilon$, it is sufficient to prove that $\sum_{xy \leqslant N} \chi_p(xy) = B_\varepsilon p^{1/2+\varepsilon}$, if N is any number from the interval $[p/(\ln^2 p), p \ln^2 p]$.

4. Elementary Consequences of a Non-elementary Theorem

Let us return to the number series (9.3.5). It forms a sequence of + or − signs. Let s be a given number, and let $p \to \infty$. We can form 2^s different sequences of + or − signs of length s, and pose the question as to whether all such sequences will be encountered in series (9.3.5) when $p \to \infty$, and if so how many times. This question was studied by Davenport in the early thirties[36], [37] who developed an analytic method for dealing with it. A number of problems of the algebraic theory of numbers cropped up at the same time and the investigation of these by Hasse led to a new trend in the algebraic and analytic theory of numbers, namely the study of Riemann's hypothesis for zeta-functions on finite fields. This trend has developed rapidly in recent years, and in 1958 led for the first time to progress in hypotheses I and I′. Before outlining these problems, we shall mention an elementary argument,

due to Brauer, regarding the sequences of length s mentioned above.

Van der Wärden (see [28]) proved the following theorem by elementary methods.

THEOREM 9.4.1. (Van der Wärden.) *Let a series of consecutive numbers, e.g. series* (9.3.11), *be split into two disjoint sets. If r is a given number, and $p > p_0(r)$, then in at least one of the two sets we can find an arithmetic progression of length r: $a\xi + b; \xi = 0, 1, \ldots, r-1$.*

We shall not give the proof of this theorem here, but instead refer the reader to Lukomskii's simple proof[28]. Our concern is with Brauer's corollary to the theorem. Let s be a given number; we choose $p > p_0(s)$ so that Van der Wärden's Theorem is applicable and take as the division of the series (9.3.11) its division into quadratic residues and nonresidues corresponding to series (9.3.5). We take the progression $a\xi + b$ ($\xi = 0, 1, \ldots, s-1$) of Theorem 9.4.1. We have

$$\chi_p(a\xi + b) = \eta_0 \qquad (\xi = 0, 1, \ldots, s-1),$$

where η_0 is the same number $+1$ or -1.

This gives us

$$\chi_p(\xi + ba') = \eta_0\chi_p(a) \quad (\xi = 0, 1, \ldots, s-1;\ aa' \equiv 1 (\mathrm{mod}\ p)).$$

The number on the right-hand side is equal to $+1$ or -1 for all values of ξ. We therefore have

THEOREM 9.4.2. *Given s and $p > p_0(s)$, the number series* (9.3.11) *always contains a sequence of not less than s consecutive quadratic residues or nonresidues.*

We shall now state a non-elementary theorem which leads to much stronger results.

Let $P_r(x) = x^r + a_1x^{r-1} + \ldots + a_r$ be an integral-valued polynomial whose coefficients are taken modulo p. We consider the "curve modulo p":

$$y^2 - P_r(x) \equiv 0 (\mathrm{mod}\ p). \tag{9.4.1}$$

Suppose that this curve is irreducible modulo p, i.e. $P_r(x) \not\equiv (Q(x))^2$ (mod p). We shall consider the integral points modulo p lying on the curve (9.4.1), i.e. those pairs of integers (x, y) for which (9.4.1) becomes an identity. We shall naturally regard as distinct pairs which are not congruent (mod p).

In 1941 Weil[51], developing the ideas of Hasse and Dering, arrived by means of very deep and non-elementary algebraic

arguments at a proof of the "Riemann hypothesis for curves modulo p". Weil's results lead to the following theorem for an irreducible curve of the form (9.4.1) (such a curve is termed elliptic for $3 \leqslant r \leqslant 4$ and hyperelliptic for $r > 4$).

THEOREM 9.4.3. (Weil.) *The number of distinct integral points modulo p on the irreducible* (mod p) *curve* (9.4.1) *is equal to*

$$p + 1 + R(p), \tag{9.4.2}$$

where

$$|R(p)| < r\sqrt{p}\,. \tag{9.4.3}$$

The following interesting corollaries to the non-elementary Theorem 9.4.3 are obtainable by elementary methods.

Let the point (x, y), taken modulo p, lie on the curve (9.4.1). In this case either $P_r(x) \equiv 0 \pmod{p}$, $y \equiv 0 \pmod{p}$, or $P_r(x) \not\equiv 0 \pmod{p}$; $\chi_p(P_r(x)) = +1$. There will be not more than r solutions (x, y) of the first type, since, by Lagrange's Theorem, the congruence $P_r(x) = 0 \pmod{p}$ has not more than r roots. For each x such that $\chi_p(P_r(\alpha)) = +1$, there are two solutions: $+y$ and $-y \pmod{p}$. We use the notation

$$S_+ = \sum_{\chi_p(P_r(x))=+1} 1\,, \qquad S_- = \sum_{\chi_p(P_r(x)=-1)} 1\,.$$

Now,

$$S_+ - S_- = \sum_{x=0}^{p-1} \chi_p(P_r(x))\,,$$

$$S_+ + S_- = p + \theta r \qquad (|\theta| \leqslant 1).$$

Hence

$$S_+ = \frac{p}{2} + \frac{1}{2} \sum_{x=0}^{p-1} \chi_p(P_r(x)) + \theta\,\frac{r}{2}\,. \tag{9.4.4}$$

In view of what was said above, we find on comparing (9.4.4) with (9.4.2) and (9.4.3) that

$$\left| \sum_{x=0}^{p-1} \chi_p(P_r(x)) \right| \leqslant 2r\sqrt{p}\,. \tag{9.4.5}$$

This deep inequality is derived by non-elementary methods, as already mentioned. Manin[18] (see Chapter 10) succeeded in 1956 in obtaining a similar inequality by elementary methods, but only for the case $r = 3$ (elliptic curves).

Let us turn to a corollary of (9.4.5). First of all, following Davenport[36], [37], we apply (9.4.5) to the problem of the distri-

bution of the above-mentioned sequences of quadratic residues of length s.†

Let $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s$ be a given sequence of signs ± 1. We form the expression

$$\Pi_s(x) = \frac{1}{2^s}(1 + \varepsilon_1\chi_p(x))(1 + \varepsilon_1\chi_p(x+1)) \cdots (1 + \varepsilon_s\chi_p(x+s-1)).$$

Obviously, $\Pi_s(x) = 1$ if the numbers $\chi_p(x), \chi_p(x+1), \ldots, \chi_p(x+s)$ form the required sequence $\varepsilon_1, \ldots, \varepsilon_s$; $\Pi_s(x) = 0$ if this is not the case. Hence

$$N_s = \sum_{x=0}^{p-1} \Pi_s(x) \tag{9.4.6}$$

furnishes the number of sequences $\varepsilon_1, \ldots, \varepsilon_s$ on the interval $[0, p-1]$ with an error not exceeding the number s; here, as is clear from (9.4.6), partial superpositions of one sequence on another are allowed.

The expression $\Pi_s(x)$ consists of the number $1/2^s$ and $2^s - 1$ expressions of the form $(1/2^s)\,\varepsilon_{\alpha_1} \ldots \varepsilon_{\alpha_t}\chi(\Phi(x))$, where $\Phi(x) = (x + \alpha_1) \ldots (x + \alpha_t)$. Given a sufficiently large p, the curves $y^2 = \Phi(x)$ are all irreducible (mod p), so that we obtain, by (9.4.5),

$$|\sum \chi_p(\Phi(x))| \leqslant 2s\sqrt{p}\,. \tag{9.4.7}$$

On taking account of this, we obtain from (9.4.6):

$$N_s = \frac{p}{2^s} + 2\theta s\sqrt{p} \qquad (|\theta| \leqslant 1). \tag{9.4.8}$$

Let

$$2^s \leqslant \frac{\sqrt{p}}{\ln^2 p}. \tag{9.4.9}$$

Now, (9.4.8) yields

$$N_s = \frac{p}{2^s} + 2\theta\frac{1}{\ln p}\frac{p}{2^s} \qquad (p > p_0). \tag{9.4.10}$$

Hence we obtain the theorem:

THEOREM 9.4.4. If $2^s \leqslant \sqrt{p}/\ln^2 p (p > p_0)$, *then there always exists a sequence* $\varepsilon_1, \ldots, \varepsilon_s$ *of signs* $+1$ or -1, *corresponding to*

† Davenport's articles[36], [37] were written before Weil had obtained his inequalities. We follow here the Leningrad University diploma dissertation of B. Z. Moroz.[21]

residues or nonresidues mod p *among the numbers* $1, 2, \ldots, p-1$. *Allowing for partial repetitions, as mentioned earlier, the number of such sequences is given by the asymptotic formula* (9.4.10).

As may be seen, this is a much stronger theorem than 9.4.2, though the latter allows of an elementary proof.

Some very remarkable corollaries to (9.4.7) were obtained by Burgess[34] in 1958. They are the first important step in the study of hypotheses I and I' of § 3 since Vinogradov's work of 1926 concerning the least quadratic nonresidue, which we mentioned in § 3. We shall state two theorems of Burgess.

THEOREM 9.4.5. (Burgess.) *No matter how small the constant* $\delta > 0$, $d(p) = B_\delta p^{(1/4)+\delta}$. *More precisely, if* $\delta > 0$, $\varepsilon > 0$ *are given constants, however small, then, given* $p > p_0(\varepsilon, \delta)$ *and any* N, *we have*

$$\left| \sum_{N=n+1}^{N+H} \left(\frac{n}{p}\right) \right| < \varepsilon H, \tag{9.4.11}$$

if $H > p^{(1/4)+\delta}$.

The next theorem follows easily from this.

THEOREM 9.4.6. (Burgess.) *For any fixed* $\delta > 0$,

$$N(p) = B_\delta p^{\frac{1}{4\sqrt{e}}+\delta}. \tag{9.4.12}$$

It will be seen that (9.4.12) is much stronger than (9.3.9); roughly speaking, the square root is extracted from the latter inequality.

In order to derive Theorem 9.4.5 from the inequality (9.4.5), some simple lemmas will be required.

LEMMA 1. *Let* r *be a positive integer,* p *a prime;* h *is an integer from the interval* $(0, p)$. *We put*

$$S_h(x) = \sum_{m=1}^{h} \left(\frac{x+m}{p}\right) \tag{9.4.13}$$

(*we put* $(s/p) = 0$ *if* $s \equiv 0 \ ((\mathrm{mod}\ p))$). *Then*

$$\sum_{x=0}^{p-1} (S_h(x))^{2r} < (2r)^r p h^r + 4r \sqrt{p} h^{2r}. \tag{9.4.14}$$

We have

$$\sum_{x=0}^{p-1} (S_h(x))^{2r} = \sum_{m_1, \ldots, m_{2r}=1}^{h} \sum_{x=0}^{p-1} \left(\frac{(x+m_1)\cdots(x+m_{2r})}{p}\right).$$

We divide the systems $m_1, \ldots, m_{2r}$ into two classes. The first class consists of the systems containing not more than r distinct

quantities, each of which is encountered an even number of times; the remaining systems form the second class. The number of systems of the first class obviously does not exceed $(2r)^r h^r$, and for each system the inner sum over x does not exceed p, so that the total sum for all these systems does not exceed $(2r)^r ph^r$. The number of systems of the second class obviously does not exceed h^{2r}, and for each the inner sum has the form

$$\sum_{x=0}^{p-1} \left(\frac{(x+n_1)^{e_1} \cdots (x+n_s)^{e_s}}{p} \right).$$

where $s \leqslant 2r$; $n_1, n_2, \ldots, n_s$ are pairwise noncongruent (mod p); $e_1, e_2, \ldots, e_s$ are not all even. We can discard the factors for which e_j is even and $x + n_j \not\equiv 0 \pmod{p}$; the number of values of x, where $x + n_j \equiv 0 \pmod{p}$, does not exceed $2r$, so that our sum differs by not more than $2r$ from the sum

$$S = \sum_x \left(\frac{(x+u_1) \cdots (x+u_k)}{p} \right),$$

where $1 \leqslant K \leqslant 2r$, $u_1, \ldots, u_k$ are pairwise noncongruent (mod p). The curve $y^2 = (x + u_1) \ldots (x + u_k)$ is irreducible (mod p), if k is odd; in this case, in accordance with (9.4.5),

$$|S| < 2u \sqrt{p} < 4r \sqrt{p}.$$

This inequality also holds when k is even. In fact, we define y from the condition

$$(x + u_1)\, y \equiv 1 (\mathrm{mod}\ p).$$

The sum S now transforms to the similar sum of $k - 1$ instead of k variables, from which we have to subtract 1, since there is no value of x corresponding to $y = 0$. In view of this, we have with even k:

$$|S| \leqslant 1 + 2(k-1) \sqrt{p} < 4r \sqrt{p}.$$

Inequality (9.4.14) can be deduced from these results.

Let us be given the integers $H > 0$, $q > 0$, t, N. We define the interval $I(q, t)$:

$$\frac{N + tp}{q} < z \leqslant \frac{N + H + tp}{q}. \tag{9.4.15}$$

LEMMA 2. *Let q run over a series of Q distinct positive integers, such that each lies in the interval*

$$(q_1, q_2) \tag{9.4.16}$$

and all are relatively prime. Let

$$2Hq_2 < p. \tag{9.4.17}$$

Now, given p, N, H, we can form for any q a series $T(q)$ of $q - Q$ integers t $(0 \leqslant t < q)$ in such a way that the intervals $I(q, t)$ do not overlap for any of the q and t in $T(q)$.

For the proof, we observe that two intervals (9.4.15) with the same q and distinct values of t are always distinct, since $0 < H < p$. Let the intervals $I(q, t)$ and $I(q', t')$ have a common point and $q > q'$. In this case

$$\frac{N + tp}{q} < \frac{N + H + t'p}{q'} \quad \text{and} \quad \frac{N + t'p}{q} < \frac{N + H + tp}{q}$$

and

$$p(tq' - t'q) + N(q' - q) < Hq,$$
$$p(tq' - t'q) + N(q' - q) > -Hq'.$$

Using (9.4.17), we find that

$$|p(tq' - t'q) + N(q' - q)| < Hq < \frac{p}{2}.$$

It follows at once from this that there exists not more than one pair t, t' for the overlapping intervals. A series $T(q)$ can now be formed for each q by eliminating from the series $0 \leqslant t < q$ all the values of t which enter into at any rate one pair t, t', corresponding to $q \neq q'$. The number of t thus eliminated does not exceed $Q - 1$, and we can therefore form a series $T(q)$ such that each includes $q - Q$ numbers t.

We turn to the proof of the theorem. It is sufficient to consider values of H under the conditions

$$p^{1/4+\delta} < H < p^{1/2+\delta}, \tag{9.4.18}$$

since the required inequality follows from (9.3.1) when $H > p^{1/2+\delta}$. Suppose that, for a certain N, and H under the conditions (9.4.18), we have

$$\left| \sum_{n=N+1}^{N+H} \left(\frac{n}{p}\right) \right| \geqslant \varepsilon H. \tag{9.4.19}$$

We shall show that this assumption leads to a contradiction for sufficiently large p. We have, for any given positive $q < p$,

$$\sum_{n=N+1}^{N+H} \left(\frac{n}{p}\right) = \sum_{t=0}^{q-1} \sum_{\substack{n=N+1 \\ n \equiv -tp \pmod q}}^{N+H} \left(\frac{n}{p}\right).$$

On putting $n = -tp + qz$ in the inner sum, we observe that z varies in the interval

$$\left[\frac{N+1+tp}{q}, \frac{N+H+tp}{q}\right]$$

and runs over the integers of the interval (9.4.15). Further,

$$\left(\frac{n}{p}\right) = \left(\frac{qz}{p}\right) = \left(\frac{q}{p}\right)\left(\frac{z}{p}\right),$$

so that we obtain from (9.4.19):

$$\sum_{t=0}^{q-1} \left| \sum_{z \in I(q,t)} \left(\frac{z}{p}\right) \right| \geqslant \varepsilon H . \tag{9.4.20}$$

We apply Lemma 2, where q runs over the primes of the interval $(\frac{1}{2}p^{1/4}, p^{1/4})$. Condition (9.4.17) is satisfied by virtue of (9.4.18). The number of integers q, which we denote by Q, is equal to

$$Q = \pi(p^{1/4}) - \pi(\tfrac{1}{2}p^{1/4}). \tag{9.4.21}$$

On summing (9.4.20) over all primes q, we find that

$$\begin{aligned} \varepsilon HQ &\leqslant \sum_q \sum_{t=0}^{q-1} \left| \sum_{z \in I(q,t)} \left(\frac{z}{p}\right) \right| \\ &\leqslant \sum_q \sum_{t \in T(q)} \left| \sum_{z \in I(q,t)} \left(\frac{z}{p}\right) \right| + \sum_q Q\, 2Hq^{-1}, \end{aligned}$$

since the number of integers z in $I(q, t)$ is less than $2Hq^{-1}$, and all except Q of the t belong to $T(q)$. Further, $\sum q^{-1} < 2p^{-1/4}Q$ by virtue of the inequalities for q. Hence we have

$$\sum_q \sum_{t \in T(q)} \left| \sum_{z \in I(q,t)} \left(\frac{z}{p}\right) \right| > QH(\varepsilon - 4p^{-1/4}Q) > \frac{\varepsilon HQ}{2}$$

for sufficiently large q (we have $Q = o(p^{1/4})$ (see Chapter 3)). Let I denote any one of our intervals $I(q, t)$. All these intervals are non-overlapping by Lemma 2; the number of them does not exceed

$$\sum_q (q - Q) < p^{1/4}Q. \tag{9.4.22}$$

We rewrite the result stated earlier in the form

$$\sum_I \left| \sum_{z \in I} \left(\frac{z}{p}\right) \right| > \frac{\varepsilon HQ}{2}. \tag{9.4.23}$$

We have, for any integer $h > 0$,

$$\sum_{z\in I}\left(\frac{z}{p}\right) = h^{-1}\sum_{m=1}^{h}\sum_{n\in I}\left(\frac{n}{p}\right)$$

$$= h^{-1}\sum_{m}^{h}\left\{\sum_{n\in I}\left(\frac{n+m}{p}\right)+\varphi_m\right\} \qquad (|\varphi_m|\leqslant 2m).$$

Hence

$$\sum_{n\in I}\sum_{m=1}^{h}\left(\frac{n+m}{p}\right) = h\sum_{z\in I}\left(\frac{z}{p}\right) - \sum_{m=1}^{h}\varphi_m.$$

Consequently,

$$\sum_{n\in I}|S_h(n)| \geqslant h\left|\sum_{z\in I}\left(\frac{z}{p}\right)\right| - 2h^2.$$

We now sum over all I and use (9.4.22) for the number of intervals I. We now obtain from (9.4.23),

$$\sum_{I}\sum_{n\in I}|S_h(n)| > \frac{\varepsilon HQh}{2} - 2p^{1/4}Qh^2.$$

We put

$$h = [\tfrac{1}{8}\varepsilon Hp^{-1/4}]. \tag{9.4.24}$$

We now obtain the inequality

$$\sum_{I}\sum_{n\in I}|S_h(n)| > \tfrac{1}{4}\varepsilon HQh.$$

We apply Hölder's inequality (see [2]). We obtain

$$\sum_{I}\sum_{n\in I}|S_h(n)| \leqslant \left\{\sum_{I}\sum_{n\in I}1\right\}^{1-1/2r}\left\{\sum_{I}\sum_{n\in I}|S_h(n)|^{2r}\right\}^{1/2r}$$

and hence

$$\sum_{I}\sum_{n\in I}|S_h(n)|^{2r} > \left(\frac{\varepsilon HQh}{4}\right)^{2r}(p^{1/4}Q\,3p^{-1/4}H)^{1-2r}$$

(we have used the fact here that the number of integers in the interval I does not exceed $3p^{-1/4}H$). Since our intervals do not overlap, we can write

$$\sum_{x}|S_h(x)|^{2r} > \left(\frac{1}{12}\varepsilon\right)^{2r}HQh^{2r}.$$

On combining this with the result of Lemma 1, we find that

$$\left(\frac{1}{12}\varepsilon\right)^{2r}HQh^{2r} < (2r)^r ph^r + 4r\sqrt{p}h^{2r}. \tag{9.4.25}$$

We have from the law of primes (see Chapter 3):

$$Q > c_1 \frac{p^{1/4}}{\ln p}.$$

We now put $r > 1/\delta$; we have $h > \varepsilon p^6/9$. Further, $H > p^{(1/4)+\delta}$. In view of this, given a sufficiently large p, the left-hand side of (9.4.25) becomes greater than the right-hand side, and we have a contradiction. This proves Theorem 9.4.5. It may easily be seen that, apart from the fundamental inequality (9.4.5), the discussion leading to Theorem 9.4.2 has here been used in a strengthened form.

Theorem 9.4.6 on the least quadratic nonresidue can be deduced from this by using the same arguments as were used for obtaining Theorem 9.3.3 from inequality (9.3.13) (see the remarks following the derivation of (9.3.23)).

CHAPTER 10

AN ELEMENTARY PROOF OF HASSE'S THEOREM

1. Statement of the Problem

Let $p > 3$ be any prime, a, b integers. We denote by N the number of solutions of the congruence

$$y^2 \equiv x^3 + ax + b \pmod{p} \tag{10.1.1}$$

(assuming that two integral-valued solutions (x_1, y_1) and (x_2, y_2) are the same if $x_1 \equiv x_2 \pmod{p}$, $y_1 \equiv y_2 \pmod{p}$). First of all, notice that the more general congruence

$$v^2 \equiv a_3u^3 + a_2u^2 + a_1u + a_0 \pmod{p}, \tag{10.1.2}$$

where $a_3 \not\equiv 0 \pmod{p}$, is easily reduced to the congruence (10.1.1). For, if we multiply both sides of (10.1.2) by a_3^2 and put $y = a_3v$, $x = a_3u + a'$, where $3a' \equiv a_2 \pmod{p}$, the new unknowns x, y will be connected by a congruence of the form (10.1.1).

If the discriminant $4a^3 + 27b^2$ of the polynomial $x^3 + ax + b$ is divisible by p, it is not difficult to find N. For in this case the congruence $x^3 + ax + b \equiv 0 \pmod{p}$ has at least a double root, so that there exist integers x_0, x_1 such that $x^3 + ax + b \not\equiv (x - x_0)^2 (x - x_1) \pmod{p}$. Hence it follows that, when $x \not\equiv x_0 \pmod{p}$, the congruence (10.1.1) is equivalent to $y^2 \equiv x - x_1 \pmod{p}$, which obviously has (with the exception of $x \equiv x_0 \pmod{p}$) p solutions if $((x_0 - x_1)/p) = -1$; $p - 2$ solutions if $((x_0 - x_1)/p) = 1$, and finally, $p - 1$ solutions if $x_0 \equiv x_1 \pmod{p}$. Corresponding to these three possible cases, we obtain $N = p + 1$, $p - 1$ or p.

It therefore remains to analyse the case when $4a^3 + 27b^2 \not\equiv 0 \pmod{p}$. This case is much more difficult, and instead of an accurate result, we can do no more than obtain an asymptotic

formula for N, with the prineipal term p and a remainder term of the order $\sqrt{p}$. To be more preeise:

THEOREM 10.1.1. *The number N of solutions of the congruence* (10.1.1) *under the condition that* $4a^3 + 27b^2 \not\equiv 0 \pmod{p}$ *satisfies the inequality*

$$|N - p| < 2\sqrt{p}. \tag{10.1.3}$$

(Evidently, we eould have ineluded the ease $4a^3 + 27b^2 \equiv 0 \pmod{p}$, inasmueh as we now have, by the above working, $|N - p| = 0$ or 1, and obviously, $1 < 2\sqrt{p}$ for any p.)

The inequality (10.1.3) is a partieular ease, first proved by Hasse[35] of the general result of Weil eoneerning the number of solutions of the eongruenee $F(x, y) \equiv 0 \pmod{p}$, where F is any polynomial in two unknowns with integral eoeffieients.

A further point: (10.1.3) is the best possible inequality, sinee it ean be shown that, when $a = -1$, $b = 0$ and $p \to \infty$, the ratio $N - p|/2\sqrt{p}$ eomes as elose as desired to unity (see [27]).

2. Addition of Solutions

Some preparatory ground-work will be needed before we ean prove the above theorem. We shall eonsider polynomials and rational funetions of an unknown x whieh is fixed once and for all time; the eoeffieients of these polynomials will be residue elasses modulo p. In aeeordanee with this understanding, we shall in future write equalities everywhere instead of eongruenees. In partieular, the equality of two polynomials $P_1(x)$ and $P_2(x)$ means that the eoeffieients of like powers of x in these polynomials represent the same residue elass (mod p), while the equality of two rational funetions $P_1(x)/Q_1(x)$ and $P_2(x)/Q_2(x)$ means the equality of the polynomials $P_1(x)\,Q_2(x)$ and $P_2(x)\,Q_1(x)$. We shall now denote by the letters a and b (as distinet from the eonvention of the previous seetion) eertain residue elasses (mod p) satisfying the relation $4a^3 + 27b^2 \not\equiv 0$.

Our purpose is primarily to estimate the number N of solutions of equation (10.1.1), though the somewhat modified equation

$$Y^2 = \frac{X^3 + aX + b}{x^3 + ax + b} \tag{10.2.1}$$

will play the most important role during the course of the proof; we shall seek the solutions (X, Y) of this latter equation in the domain of rational functions of the unknown x (with coefficients of the type described).

Let us look more closely at (10.2.1). First of all, it has the obvious solution $X = x$, $Y = \pm 1$. A less obvious solution which is extremely important for what follows is $X = x^p$, $Y = \pm (x^3 + ax + b)^{1/2(p-1)}$ (it is sufficient to observe that $(x^3 + ax + b)^p = x^{3p} + ax^p + b$, because the binomial coefficients $p!/i!\, j!\, k!$ $(i + j + k = p)$ are divisible by p when $ijk \not\equiv 0 \pmod{p}$.

A simple method exists for finding a third solution of (10.2.1) if two solutions are already known. This method is best described in geometric terms; the equation (10.2.1) is regarded as a "curve" with the current coordinates X, Y (the unknown x plays the role of parameter, on which the coefficients of the equation of the curve depend). Let us consider a pair of points (X_1, Y_1), (X_2, Y_2) on the curve; through these we draw a straight line with current coordinates X, Y:

$$\frac{X - X_1}{X_2 - X_1} = \frac{Y - Y_1}{Y_2 - Y_1}. \tag{10.2.2}$$

In order for equation (10.2.2) to have a meaning, we assume temporarily that $X_2 \neq X_1$ (i.e. the chosen points are not coincident). The straight line (10.2.2) cuts the curve (10.2.1) in three points, two of which—(X_1, Y_1) and (X_2, Y_2)—are known. To find the third point, we express Y in terms of X from (10.2.2) and substitute the result in (10.2.1):

$$X^3 + aX + b - \left(\frac{Y_2 - Y_1}{X_2 - X_1}(X - X_1) + Y_1\right)^2 (x^3 + ax + b) = 0.$$

The two roots X_1 and X_2 of this cubic equation in X are known; the coefficient of X^2 on the left-hand side is the sum of all three roots with reversed sign and is equal to

$$-\left(\frac{Y_2 - Y_1}{X_2 - X_1}\right)^2 (x^3 + ax + b).$$

Hence the third root is

$$X_3 = -X_1 - X_2 + \left(\frac{Y_2 - Y_1}{X_2 - X_1}\right)^2 (x^3 + ax + b). \tag{10.2.3}$$

On substituting this value in the equation of the straight line, we find the second coordinate of the required point of intersection:

$$Y_3 = \frac{Y_2 - Y_1}{X_2 - X_1}(X_3 - X_1) + Y_1. \tag{10.2.4}$$

In addition to the point (X_3, Y_3), the point $(X_3, -Y_3)$ symmetric to it also lies on the curve (10.2.1), and it is this that we shall regard as the final result of the operation just described on the points (X_1, Y_1) and (X_2, Y_2).

This operation is denoted by the symbol $+$ and is termed "addition" of the points of the curve (10.2.1) or, what amounts to the same thing, of the solutions of equation (10.2.1). We therefore write

$$(X_1, Y_1) + (X_2, Y_2) = (X_3, -Y_3), \tag{10.2.5}$$

if X_3, Y_3 are expressible in terms of the coordinates of the two given points in accordance with (10.2.3) and (10.2.4).

It still remains to consider the degenerate case $X_1 = X_2$. If at the same time $Y_1 \neq Y_2$, the straight line $X - X_1 = 0$ passing through the points (X_1, Y_1) and $(X_1, -Y_1)$ cuts the curve (10.2.1) at a further, infinitely remote, ideal point, which coincides with the point symmetrical to it. We are not going to assign a strict meaning to this phrase: in future we shall get round the possibility of such a "solution" of equation (10.2.1) appearing in our construction by means of a formal device.

Finally, if $X_1 = X_2$, $Y_1 = Y_2$, then, as the geometrical interpretation of (10.2.3) and (10.2.4) would suggest, instead of (10.2.2) we have to write down the equation of the "tangent" to the curve at the point (X_1, Y_1):

$$Y - Y_1 = \frac{3X_1^2 + a}{2Y_1(x^3 + ax + b)}(X - X_1), \tag{10.2.6}$$

then solve this equation simultaneously with the equation (10.2.1). We finally obtain

$$X_3 = -2X_1 + \frac{(3X_1^2 + a)^2}{4(X_1^3 + aX_1 + b)}. \tag{10.2.7}$$

The second coordinate Y_3 of the point of intersection is obtained if we substitute this value for X in (10.2.1); we recall once more that, to obtain the double point (X_1, Y_1), we have to take the symmetric point $(X_3, -Y_3)$ instead of the point (X_3, Y_3) itself.

The addition of points of the curve (10.2.1), as defined by (10.2.3); (10.2.4); (10.2.6); (10.2.7), satisfies all the axioms of a group operation. It is commutative: this follows directly from the definition. It also has a zero—this is the infinitely remote point. Finally, it is associative. Since we shall never require these last two properties, we shall not dwell on their proof, all the more, because direct verification of the associative law from the addition formulae leads to fairly unwieldy working. We shall only have occasion below to use the following fact: equation (10.2.5) is equivalent to the equation

$$(X_1, Y_1) + (X_3, Y_3) = (X_2, -Y_2). \tag{10.2.8}$$

The geometric interpretation suggests how this assertion can be verified without any working: the same straight line passes through the points (X_1, Y_1) and (X_3, Y_3) as passes through (X_1, Y_1) and (X_2, Y_2), while the role of the third point of intersection is now played by (X_2, Y_2).

3. The Fundamental Construction

We shall now stop talking about points of a curve and again describe as a solution of equation (10.2.1) the pair (X, Y) of rational functions of the unknown x satisfying this equation. We indicated two particular solutions at the start of the last section. We now use the operation of addition just defined to construct from these two solutions a sequence of solutions, infinite in both directions, which represents so to speak an arithmetic progression with the difference $(x, 1)$ and the initial term $(x^p, (x^3 + ax + b)^{1/2(p-1)})$. We obtain, in fact,

$$(X_0, Y_0) = (x^p, (x^3 + ax + b)^{p-1/2}) \tag{10.3.1}$$

and, if the solution (X_n, Y_n) is already defined,

$$(X_{n+1}, Y_{n+1}) = (X_n, Y_n) + (x, 1), \tag{10.3.2}$$

$$(X_{n-1}, Y_{n-1}) = (X_n, Y_n) + (x, -1). \tag{10.3.3}$$

(It is clear that formula (10.3.2) is to be used for continuing the sequence rightwards from any point, and (10.3.3) for continuing it leftwards.) If it turns out during the construction that $(X_n, Y_n) = (x, -1)$ for a certain n, then we leave a gap instead of the solution (X_{n+1}, Y_{n+1}) and put $(X_{n+2}, Y_{n+2}) = (x, 1)$. Similarly, if we

arrive at the solution $(X_n, Y_n) = (x, 1)$ when continuing leftwards, then we leave a gap instead of the solution (X_{n-1}, Y_{n-1}) and put $(X_{n-2}, Y_{n-2}) = (x, -1)$. Obviously, these conventions uniquely determine a sequence of solutions of equation (10.2.1), possibly containing gaps.

Given any n for which the solution (X_n, Y_n) is defined, we write X_n as an irreducible ratio P_n/Q_n of two polynomials of the unknown x. We define this ratio uniquely by requiring that the leading coefficient of P_n be unity.

We finally introduce a sequence of non-negative integers, infinite in both directions, by putting

$$d_n = \begin{cases} 0 \text{ if the function } X_n \text{ is not defined,} \\ \text{the degree of the polynomial } P_n \text{ otherwise.} \end{cases}$$

We are now in a position to explain the connection between the construction described above with the problem that we posed originally, of estimating the number N of solutions of the congruence (10.1.1). The first step in our explanation is

LEMMA 1.

$$d_{-1} - d_0 - 1 = N_p - p.$$

We have, by (10.2.3):

$$X_{-1} = -x - x^p + \frac{[1 + (x^3 + ax + b)^{p-1/2}]^2 (x^3 + ax + b)}{(x - x^p)^2}$$

Notice that the degree of the numerator P_{-1} is unity greater than the degree of the denominator Q_{-1}. For, it is easily seen that

$$X_{-1} = \frac{x^{2p+1} + R(x)}{(x - x^p)^2},$$

where $R(x)$ is a polynomial of degree not exceeding $2p$.

We calculate the degree of the denominator Q_{-1} by carrying out the necessary cancellings. On the one hand, $(x - x^p)^2 = [x(x-1) \dots (x - p + 1)]^2$; on the other hand, as we know,

$$(x^3 + ax + b)^{p-1/2} = \left(\frac{x^3 + ax + b}{p}\right).$$

Thus, when cancelling the denominator, the only factors that fall out are of the form $(x - k)^2$ with $(k^3 + ak + b)/p = -1$ and of the form $x - l$ with $l^3 + al + b = 0$. Let m be the number of

factors of the first type, and n the number of the second. Now, $d_{-1} = 2p - 2m - n + 1$, $d_{-1} - d_0 - 1 = p - 2m - n$, because $d_0 = p$. The number N is equal to $2(p - m) - n$, because for each class of residues k with $(k^3 + ak + b)/p = 1$ there are two corresponding solutions of the congruence (10.1.1), while for each class of residues l, for which $l^3 + al + b = 0$, there is only one corresponding solution $(l, 0)$. This proves what we require.

The fundamental difficulties of the proof of (10.1.3) are concentrated in the proof of the following.

FUNDAMENTAL LEMMA.

$$d_{n-1} + d_{n+1} = 2d_n + 2.$$

4. Deduction of the Theorem from the Fundamental Lemma

Let us show how (10.1.3) is derived from the fundamental lemma.

LEMMA 2.

$$d_n = n^2 - (d_{-1} - d_0 - 1)\, n + d_0.$$

In accordance with the fundamental lemma, $d_{n+2} = 2d_{n+1} - d_n + 2$.

Obviously, the lemma is true for $n = -1$ and $n = 0$; let it be true for the indices n and $n + 1$, $n \geqslant 0$. Then

$$\begin{aligned} d_{n+2} = 2d_{n+1} - d_n + 2 &= 2[(n+1)^2 - (d_{-1} - d_0 - 1)(n+1) + d_0] \\ &\quad - n^2 + (d_{-1} - d_0 - 1)\, n - d_0 + 2 \\ &= (n+2)^2 - (d_{-1} - d_0 - 1)(n+2) + d_0. \end{aligned}$$

We can carry out the induction for $n < 0$ in the same way.

Proof of (10.1.3). From Lemmas 1 and 2, $d_n = n^2 - (N - p)\, n + p$. This quadratic expression takes non-negative values for all integers n, while it cannot have two consecutive integral zeros by definition of the numbers d_n. Let us show that the discriminant of such an expression is non-positive. If this were not the case, the expression would have two roots α and β in the interval between two consecutive integers: $n \leqslant \alpha < \beta \leqslant n + 1$, where the signs of equality cannot both hold simultaneously. But in this case the numbers $\alpha\beta$ and $\alpha + \beta$ could not simultaneously be integers. Consequently,

$$(N - p)^2 - 4p \leqslant 0,$$

so that

$$|N - p| < 2\sqrt{p}.$$

The theorem is proved.

5. Proof of the Fundamental Lemma

We proceed to the proof of the lemma. We shall write deg P for the degree of the polynomial P in the unknown x. The following auxiliary lemma will be needed.

LEMMA 3. *For all n, for which the function X_n is defined, we have*

$$\deg P_n > \deg Q_n.$$

We shall prove this inequality by finding formally the value of the function X_n for $x = \infty$. Let m be zero or the first number after the recurring gap, $m \geqslant 0$. By construction, $X_m|_\infty = \infty$, $Y_m|_\infty \neq 0$. Assuming that these relations hold for some index n, following which there is no gap, we can show that they hold for the index $n + 1$. This is sufficient for the assertion to hold concerning the degrees of the numerator and denominator of the function X_{n+1}. As (10.2.1) indicates, it is sufficient to verify merely that $Y_{n+1}|_\infty \neq 0$. Suppose the converse. In view of the fact that the fraction

$$\frac{X_{n+1}^3 + aX_{n+1} + b}{x^3 + ax + b}$$

must be a square, the difference between the degrees of the numerator and denominator of X_{n+1} must be an odd number, which, in conjunction with $Y_{n+1}|_\infty = 0$, yields $X_{n+1}|_\infty = 0$. It follows from (10.2.4) and (10.3.2) that

$$Y_{n+1} = \frac{1 - Y_n}{x - X_n}(x - X_{n+1}) - 1.$$

Hence we find that

$$\frac{1 - Y_n}{x - X_n}(x - X_{n+1})\Bigg|_\infty = 1 \quad \text{or} \quad \frac{1 - Y_n}{1 - \dfrac{X_n}{x}}\Bigg|_\infty = 1,$$

i.e.

$$\frac{Y_n x}{X_n}\Bigg|_\infty = 1, \quad \frac{Y_n^2 x^2}{X_n^2}\Bigg|_\infty = \frac{(X_n^3 + aX_n + b)\,x^2}{(x^3 + ax + b)\,X_n^2}\Bigg|_\infty = 1.$$

Since $X_n|_\infty = x|_\infty = \infty$, it follows from this that $X_n/x|_\infty = 1$. On the other hand, it follows from (10.2.3) and (10.3.2) that

$$X_{n+1} = -x - X_n + \frac{(1 - Y_n)^2}{(x - X_n)^2}(x^3 + ax + b). \qquad (10.5.1)$$

Hence we find that

$$\frac{X_{n+1}}{x} = -1 - \frac{X_n}{x} + \frac{(1 - Y_n)^2}{\left(1 - \frac{X_n}{x}\right)^2}\left(1 + \frac{a}{x^2} + \frac{b}{x^3}\right),$$

so that

$$\left.\frac{X_{n+1}}{x}\right|_\infty = -1.$$

But it follows from this equation, in view of (10.2.1), that $Y_{n+1}|_\infty = -1$, and this contradicts our assumption that $Y_{n+1}|_\infty = 0$.

The leftwards induction proceeds in the same way.

The lemma is proved.

During the proof of the fundamental lemma we shall distinguish two cases.

(a) Not all the functions X_{n-1}, X_n, X_{n+1} are defined. It is clear that only one of the set of three consecutive functions needs to be undefined. If the middle function is not defined, then

$$X_{n-1} = x, \quad X_{n+1} = x, \quad d_{n-1} = 1, \quad d_n = 0, \quad d_{n+1} = 1,$$

which satisfies the equation we wish to prove. If one of the extreme functions, say the one on the right, is not defined, then by virtue of (10.2.7):

$$X_n = x, \quad X_{n+1} = \frac{(x^2 - a)^2 - 8bx}{4(x^3 + ax + b)},$$

i.e. $d_n = 1$, $d_{n+1} = 4$, which in conjunction with $d_{n-1} = 0$ again yields the assertion of the fundamental lemma. The case when the left-hand function is not defined may be treated similarly.

(b) All the functions X_{n-1}, X_n, X_{n+1} are defined.

On reducing to a common denominator and collecting like terms in (10.2.8), we find that

$$\begin{aligned} X_{n-1} &= \frac{-(xQ_n + P_n)(xQ_n - P_n)^2 + (1 + Y_n)^2(x^3 + ax + b)Q_n^2}{(xQ_n - P_n)^2} \\ &= \frac{(xQ_n + P_n)(xP_n + aQ_n) + 2bQ_n^2 - 2Y_n(x^3 + ax + b)Q_n^2}{(xQ_n - P_n)^2} \\ &= \frac{R}{(xQ_n - P_n)^2}. \end{aligned} \qquad (10.5.2)$$

Similarly,

$$X_{n+1} = \frac{-(xQ_n + P_n)(xQ_n - P_n)^2 + (1 - Y_n)^2(x^3 + ax + b)Q_n^2}{(xQ_n - P_n)^2}$$

$$= \frac{(xQ_n + P_n)(xP_n + aQ_n) + 2bQ_n^2 + 2Y_n(x^3 + ax + b)Q_n^2}{(xQ_n - P_n)^2}$$

$$= \frac{S}{(xQ_n - P_n)^2}. \tag{10.5.3}$$

On cross-multiplying (10.5.2) and (10.5.3) and cancelling, we obtain (we shall omit the details of the working)

$$X_{n-1}X_{n+1} = \frac{P_{n-1}P_{n+1}}{Q_{n-1}Q_{n+1}} = \frac{(xP_n - aQ_n)^2 - 4bQ_n(xQ_n + P_n)}{(xQ_n - P_n)^2}. \tag{10.5.4}$$

The aim of our next arguments is to show that

$$Q_{n-1}Q_{n+1} = (xQ_n - P_n)^2. \tag{10.5.5}$$

The fundamental lemma follows quite speedily from this equality. For, it now follows from (10.5.5) that

$$P_{n-1}P_{n+1} = (xP_n - aQ_n)^2 - 4bQ_n(xQ_n + P_n)$$

i.e.

$$d_{n-1} + d_{n+1} = \deg.(P_{n-1}P_{n+1}) = \deg.(x^2P_n^2) = 2d_n + 2,$$

since, by Lemma 3, the leading term of the polynomial $P_{n-1}P_{n+1}$ is the same as the leading term of x^2P^2.

It remains to prove (10.5.5).

We recall that, in the domain of polynomials, there is a unique arrangement into irreducible factors. Let E be an irreducible polynomial, $e > 0$ any integer. We shall say that the polynomial E^e divides exactly into an irreducible rational function if its numerator is divisible by E^e but is not divisible by E^{e+1}.

To prove (10.5.5), it is sufficient to show that, if the polynomial E^e divides exactly into $(xQ_n - P_n)^2$, then it also divides exactly into $Q_{n-1}Q_{n+1}$. For, in this case the quotient $Q_{n-1}Q_{n+1}/(xQ_n - P)^2$ is a polynomial which is relatively prime to the polynomial $(xQ_n - P_n)^2$. But, since it follows from (10.2.1) that the function $Y_n(x^3 + ax + b)Q_n^2$ is a polynomial, it easily follows from (10.5.2) and (10.5.3) that the denominators Q_{n-1} and Q_{n+1} divide into the polynomial $(xQ_n - P_n)^4$. Thus the quotient $Q_{n-1}Q_{n+1}/(xQ_n - P_n)^2$ must be a constant, and this constant is unity by virtue of our normalization of the leading terms of the numerators P_n.

We split all the irreducible divisors E of the polynomial $(xQ_n - P_n)^2$ into three groups. (Notice that all such polynomials E divide into RS.)

We put in the first group the polynomials E which divide into R but do not divide into S. It follows at once from (10.5.2)–(10.5.4) that, if the polynomial E^e divides exactly into $(xQ_n - P_n)^2$, then it divides exactly into the denominator Q_{n+1} and is relatively prime to the denominator Q_{n-1}.

We put in the second group the polynomials E which divide into S but not into R. We find similarly that, if the polynomial E^e divides exactly into $(xQ_n - P_n)^2$, then it divides exactly into Q_{n-1} and is relatively prime to Q_{n+1}.

Finally, we put into the third group the polynomials E which divide into both R and S. Since

$$P_n \equiv xQ_n \pmod{E},$$

it follows from (10.5.2) and (10.5.3) that

$$\begin{aligned} R &\equiv 2Q_n^2(1 - Y_n)\,(x^3 + ax + b) \pmod{E}, \\ S &\equiv 2Q_n^2(1 + Y_n)\,(x^3 + ax + b) \pmod{E}. \end{aligned} \tag{10.5.6}$$

A polynomial E which divides into the polynomial $(xQ_n - P_n)^2$ cannot divide into Q_n, since P_n and Q_n are relatively prime. It follows from (10.5.6) that $S + R \equiv 4Q_n^2(x^3 + ax + b) \pmod{E}$, so that, if E divides into R and S, it divides exactly into the polynomial $x^3 + ax + b$ (on the assumption that this latter has no repeated zeros).

Thus, let E be an irreducible divisor of the polynomial $x^3 + ax + b$. We assume at first that $Y_n \not\equiv \pm 1 \pmod{E}$ (by definition, this notation means that the numerator of the irreducible representation of the function $Y_n \pm 1$ is not divisible by E). It now follows from (10.5.2) that E divides exactly into R, since the polynomial $(xQ_n - P_n)^2$ is divisible at least by E^2. Similarly, we can show that E divides exactly into S; but in this case it follows from (10.5.2)–(10.5.4) that E^2 divides exactly into $(xQ_n - P_n)^2$.

It therefore remains to consider the case $Y_n \equiv \pm 1 \pmod{E}$. Let say $Y_n \equiv -1 \pmod{E}$ (the second possibility may be considered in the same way). Now E divides exactly into S. Let E^{2e} divide exactly into $(xQ_n - P_n)^2$, while E^{2f+1} divides exactly into $(1 + Y_n)^2\,(x^3 + ax + b)\,Q^2$. Obviously, E^{2f} divides exactly into

the function $(1 + Y_n)^2 (1 - Y_n)^2 = (1 - Y^2)^2$. But

$$(1 - Y_n^2)^2 = \frac{(x^3 + ax - X_n^3 - a X_n)^2}{(x^3 + ax + b)^2}$$
$$= \frac{(x - X_n)^2 (x^2 + xX_n + X_n^2 + a)^2}{(x^3 + ax + b)^2}.$$

In addition, $x \equiv X_n \pmod{E}$ and $x^2 + xX_n + X^2 + a \equiv 3x^2 + a \not\equiv 0 \pmod{E}$, so that $2f = 2e - 2$, and consequently the number $2f + 1 = 2e - 1$ is less than the degree to which E divides exactly into $(xQ_n + P_n)(xQ_n - P_n)^2$. Hence E^{2e-1} divides exactly into R, while E^{2e} divides exactly into RS. Now, it again follows from (10.5.2)–(10.5.4) that E^2 divides exactly into $Q_{n-1}Q_{n+1}$.

This completes the proof of the fundamental lemma and the formula (10.1.3).

CHAPTER 11

AN ELEMENTARY PROOF OF SIEGEL'S THEOREM

1. Statement of the Theorem. Method of Proof

Siegel's theorem on the number $h(\Delta)$ of classes of binary quadratic forms of the fundamental discriminant Δ asserts that if $\Delta = -D < 0$, then

$$\lim_{D\to\infty} \frac{\ln h(-D)}{\ln D} = \frac{1}{2}. \tag{11.1.1}$$

This theorem, proved by Siegel in 1935[48] by analytic methods, plays a basic role in modern analytic theory of numbers. We shall here give a proof of the theorem using fairly elementary methods; this proof, due to Linnik,[15] is based on an idea of Vinogradov.

In 1918 Vinogradov pointed out the remarkable similarity of the behaviour of the primitive real characters $\chi(n)$ modulo large D in the interval $[1, D-1]$ to the behaviour of the Liouville function $\lambda(n) = (-1)^{\nu(n)}$ ($\nu(n)$ is the number of the prime divisors of n) in the same interval. The basis of this similarity lies in the complete multiplicativeness of both types of function. But, whereas Vinogradov's formula (see [3])

$$\sum_{n\leqslant x} \chi(n) = B\sqrt{D}\ln D \tag{11.1.2}$$

holds for the sums of characters $\chi(n)$, for any $x \in [1, D-1]$, the existence of a formula of the type

$$\sum_{n\leqslant x} \lambda(n) = Bx^{1-c_0} \qquad (c_0 > 0 \text{ is a constant}) \tag{11.1.3}$$

for the Liouville function $\lambda(n)$ constitutes an unsolved problem, similar to the "weak Riemann hypothesis" for the zeta functions of Riemann.

To explain the connection between the formulae (11.1.1) and (11.1.2), Vinogradov suggested that it would be useful to study the amount of "damage" done if the Liouville function $\lambda(n)$ is replaced by the character $\chi(n)$. In devising an elementary proof of Siegel's theorem on the basis of this idea, we shall utilize the following:

(a) the elements of the arithmetic theory of binary quadratic forms and real characters, including Vinogradov's formula (11.1.2);†
(b) the elementary properties of primes and the Möbius function;
(c) the definition and properties of the base e of natural logarithms;
(d) elementary algebra;
(e) the concept of limit (this can be avoided by making an obvious modification to the statement of the theorem), and an elementary property of a continuous function.

Unfortunately, the proof is just as ineffective as the other (familiar) proofs.

The formula (11.1.1) does not provide a basis for effective calculations.

2. Lemmas

As in earlier chapters,

$$C_0, C_1, \ldots;\ c_0, c_1, \ldots;\ \varepsilon_0, \varepsilon_1, \ldots;\ \eta_0, \eta_1, \ldots,$$

denote positive constants; in the last two series the constants are always small (less than 1/4).

Given the real primitive character $\chi_k(n)$ modulo $k > 2$ and the number $\beta > \frac{1}{2}$, not exceeding 1.

We shall say that $\chi_k(n)$ has the property $\mathfrak{A}(\beta)$ if there exists a constant $C_0(\chi_k)$ such that, for any $N > C_0(\chi_k)$ in the interval $[\sqrt{N}, N]$, there is an $N_1 \in [\sqrt{N}, N]$ such that

$$\left| \sum_{n \leqslant N_1} \chi_k(n)\, \mu(n) \right| > N_1^{\beta}. \tag{11.2.1}$$

The following lemma is concerned with the property $\mathfrak{A}(\beta)$.

† By the way, this can be replaced by a trivial formula with D instead of $\sqrt{D} \ln D$.

FIRST FUNDAMENTAL LEMMA. *If there exists a character* $\chi_k(n)$ *possessing the property* $\mathfrak{N}(\beta)$, *where* $\beta \geqslant 3/4$, *then, when* $D > C_1(\chi_k, \eta_0)$, *we have*

$$h(-D) > D^{1/2-\eta(\beta)}, \tag{11.2.2}$$

where

$$\eta(\beta) = 10{\cdot}5(1-\beta) + \eta_0 \tag{11.2.3}$$

and η_0 *is a fixed positive number, as small as desired.*

The proof requires several auxiliary lemmas.

LEMMA 1. *If* $\varrho > 1$, $q_1 > q_0 > 1$, *then*

$$\sum_{q=q_0}^{q_1} \frac{1}{q^\varrho} < \frac{2}{1-\frac{1}{2^\varrho}} \frac{1}{q_0^{\varrho-1}}. \tag{11.2.4}$$

To prove this, we replace all the numbers q between q_0 and $2q_0$ by q_0, all the numbers q between $2q_0 + 1$ and $4q_0$ by $2q_0$, and so on, then consider the sum of the geometric progression with ratio $\frac{1}{2^\varrho}$.

LEMMA 2. *If*

$$\left| \sum_{n \leqslant N_1} \chi_k(n)\, \mu(n) \right| > N_1^\beta \left(\beta > \frac{1}{2}\right) \tag{11.2.5}$$

and ε *is any number under the condition* $0 < \varepsilon < \beta - \frac{1}{2}$, *then, when* $N_1 > C_0(\beta, \varepsilon, \chi_k)$, *there exists in the interval* $[N_1^{2\beta-1-2\varepsilon}, N_1]$ *an integer* $N_2 \in [N_1^{2\beta-1-2\varepsilon}, N_1]$ *such that*

$$\left| \sum_{n \leqslant N_2} \chi_k(n)\, \lambda(n) \right| > N_2^{\beta-\varepsilon}. \tag{11.2.6}$$

Using the "sieve" method, we have

$$\sum_{n \leqslant N_1} \chi_k(n)\, \mu(n) = \sum_{q \leqslant \sqrt{N_1}} \mu(q) \sum_{m \leqslant N_1/q_2} \chi_k(m)\, \lambda(m). \tag{11.2.7}$$

We pick out the values of q for which $q \geqslant N_1^{1-\beta+\varepsilon}$. We obtain for these, in accordance with (11.2.4):

$$\left| \sum_{N_1^{1-\beta+\varepsilon} < q \leqslant \sqrt{N_1}} \mu(q) \sum_{m \leqslant N_1/q^2} \chi_k(m)\, \lambda(m) \right| \leqslant \sum_{N_1^{1-\beta+\varepsilon} < q \leqslant \sqrt{N_1}}' \frac{N_1}{p^2} < \frac{4N_1}{N_1^{1-\beta+\varepsilon}}.$$

This last number does not exceed $4N_1^{\beta-\varepsilon}$, so that

$$\sum_{n \leqslant N_1} \chi_k(n)\, \mu(n) = \sum_{q < N_1^{1-\beta+\varepsilon}} \mu(q) \sum_{m \leqslant N_1/q^2} \chi_k(m)\, \lambda(m) + R_{N_1}, \tag{11.2.8}$$

where $|R_{N_1}| < 4N^{\beta-\varepsilon}$.

Suppose now that, when $q < N_1^{1-\beta+\varepsilon}$, we always have

$$\left| \sum_{n \leqslant N_1/q^2} \chi_k(m)\, \lambda(m) \right| < \left(\frac{N_1}{q^2}\right)^{\beta-\varepsilon},$$

and let us show that a contradiction is implied. Using (11.2.4), we obtain from our assumption:

$$\left| \sum_{q<N_1^{1-\beta+\varepsilon}} \mu(q) \sum_{m \leqslant N_1/q^2} \chi_k(m)\, \lambda(m) \right|$$

$$< N_1^{\beta-\varepsilon} \sum_{1 \leqslant q \leqslant N_1^{1-\beta+\varepsilon}} \frac{1}{q^{2\beta-2\varepsilon}} < C(\beta, \varepsilon)\, N_1^{\beta-\varepsilon}.$$

On substituting in (11.2.8), we find that

$$\left| \sum_{n \leqslant N_1} \chi_k(n)\, \mu(n) \right| < (C(\beta, \varepsilon) + 4)\, N_1^{\beta-\varepsilon},$$

which contradicts (11.2.5). Hence it follows that, given a sufficiently large N_1, there exists q such that

$$\left| \sum_{m \leqslant N_1/q^2} \chi_k(m)\, \lambda(m) \right| \geqslant \left(\frac{N_1}{q^2}\right)^{\beta-\varepsilon}, \qquad q < N_1^{1-\beta+\varepsilon}.$$

Putting $N_2 = N_1/q^2$, we see that $N_2 \leqslant N_1$ and $N_2 > N^{2\beta-1-2\varepsilon}$. This proves Lemma 2.

LEMMA 3. *We have for any character $\chi(n)$, primitive* (mod D),

$$\sum_{n \leqslant x} \chi(n) = B\sqrt{D}\, \ln D, \tag{11.2.9}$$

where x is any number.

Further, if $\nu(n)$ is the number of prime divisors of n, $\tau(n)$ is the number of all its divisors, then

$$2^{\nu(n)} = B_\varepsilon n^\varepsilon; \quad \tau(n) = B_\varepsilon n^\varepsilon$$

for any fixed ε and $n \to \infty$.

These three relations, of which (11.2.9) is the familiar Vinogradov formula, may be proved by elementary methods.

LEMMA 4. *If $Q(x, y) = ax^2 + bxy + cy^2$ is a reduced positive quadratic form with discriminant $-D$, then*

$$\sum_{m \leqslant M} r(m) < c_1 \frac{M}{\sqrt{D}},$$

where $r(m)$ is the number of solutions of the equation $Q(x, y) = m$, and $M \geqslant D^2$.

Here, it is sufficient to make a rough estimate of the number of integral points in the region $Q(x, y) \leqslant M$.

3. Proof of the First Fundamental Lemma

Given the real character $\chi_k(n)$ possessing the property $\mathfrak{P}(\beta)$ with $\beta \geqslant 3/4$, and the corresponding constant $C_0(\chi_k)$. Now, by definition of the $\mathfrak{P}(\beta)$ property and Lemma 2, given any positive $\varepsilon < \beta - \frac{1}{2}$ and $N > C_1(\beta, \varepsilon, \chi_k)$ in the interval $[N^{\beta-1/2-\varepsilon}, N]$, there exists a number N_2 such that

$$\left| \sum_{n \leqslant N_2} \chi_k(n)\, \lambda(n) \right| > N_2^{\beta-\varepsilon}. \tag{11.3.1}$$

We put here:

$$\varepsilon = \varepsilon_0 = 0{\cdot}01(1 - \beta). \tag{11.3.2}$$

Suppose there exists an infinite sequence of fundamental discriminants $-D_j$ $(j = 1, 2, \ldots)$ under the conditions

$$h(-D_j) \leqslant D_j^{1/2-\eta(\beta)}; \quad \eta(\beta) = 10{\cdot}5(1-\beta) + \eta_0, \tag{11.3.3}$$

where η_0 is a fixed positive number, as small as desired. We have to show that this assumption implies a contradiction. We take $D_j = D$ so large that $D_j > C_1(\beta, \varepsilon, \chi_k)$, and choose

$$N = \left[2D^{\frac{4}{\beta-\frac{1}{2}-\varepsilon_0}} \right] = N_3.$$

There now exists $N_2 \in [D^4, N_3]$, $N_3 > D^8$, such that

$$\left| \sum_{n \leqslant N_2} \chi_k(n)\, \lambda(n) \right| > N_2^{\beta-\varepsilon_0}. \tag{11.3.4}$$

We put $\chi(n) = (-D/n)$ and introduce the sum

$$\sum_{n \leqslant N_2} \chi_k(n)\, \chi(n).$$

By Lemma 3 and the fundamental properties of characters, we have

$$\sum_{n \leqslant N_2} \chi_k(n)\, \chi(n) \ll \sqrt{kD}\, \ln(kD). \tag{11.3.5}$$

During our investigation of the possible differences in the behaviour of $\lambda(n)$ and $\chi(n)$, we shall compare the sums (11.3.4) and (11.3.5).

First let $p_1, p_2, \ldots, p_s$ be distinct primes for which $\chi(p_j) = 0$. Now, $p_1 p_2 \ldots p_s \leqslant D$.

We introduce the multiplicative function $\chi'(n)$, which differs from $\chi(n)$ solely in the fact that $\chi'(p_j) = -1$ $(j = 1, 2, \ldots, s)$. We now obtain:

$$\sum_{n \leqslant N_2} \chi_k(n)\,\chi'(n) = \sum_{n \leqslant N_2} \chi_k(n)\,\chi(n) - \sum_{p_j} \sum_{n \leqslant N_2/p_j} \chi_k(n)\,\chi(n) + \cdots + (-1)^{s-1} \sum_{n \leqslant \frac{N_2}{p_1 \cdots p_s}} \chi_k(n)\,\chi(n).$$

The number of our sums does not exceed $2^s \leqslant \tau(D) \leqslant B_{\varepsilon'} D^{\varepsilon'}$ whatever the fixed $\varepsilon' > 0$. By Lemma 3, each of the sums does not exceed $C_1 \sqrt{Dk} \ln (kD)$, so that

$$\sum_{n \leqslant N_2} \chi_k(n)\,\chi'(n) = B_{\varepsilon'} N_2^{\varepsilon'} \sqrt{kD} \ln (kD). \tag{11.3.6}$$

We replace all the $\chi'(n)$ in this sum by $\lambda(n)$, which is equivalent to replacing the values $\chi'(p) = 1$ for $p|n$ by $\lambda(p) = -1 (p|n)$.

We denote by $\mathfrak{N}_D$ the set of all such numbers n', each prime divisor $p|n'$ of which will have $\chi(p) = +1$.

It is a well known fact from the elementary theory of quadratic forms that $\mathfrak{N}_D$ coincides with the set of all the numbers primitively representable by forms with the discriminant $-D$. Consequently,

$$\sum_{n \leqslant N_2} \chi_k(n)\,\lambda(n) = \sum_{n \leqslant N_2} \chi_k(n)\,\chi'(n) - 2 \sum_{p_j \in \mathfrak{N}_D} \sum_{n \leqslant N_2/p_j} \chi_k(n)\,\chi'(n) + 4 \sum_{p_i p_j \in \mathfrak{N}_D} \sum_{n \leqslant N_2/p_i p_j} \chi_k(n)\,\chi'(n) + \cdots + (-1)^m 2^m \sum_{p_i p_j \cdots p_\varrho \in \mathfrak{N}_D} \sum_{n \leqslant \frac{N_2}{p_i p_j \cdots p_\varrho}} \chi_k(n)\,\chi'(n) + \cdots \tag{11.3.7}$$

We divide the sums into two types:

(1) those in which $p_i p_j \ldots p_\varrho \geqslant N_2/D^2$,

(2) the rest.

The sums of the first type jointly do not exceed

$$N_2 \sum_{N_2/D^2 \leqslant n' \leqslant N_2} \frac{2^{\nu(n')}}{n'},$$

where n' is a square-free number belonging to $\mathfrak{N}_D$. It follows from the elementary theory of quadratic forms that this sum does not exceed

$$N_2 \sum_{Q} \sum_{N_2/D^2 \leqslant Q(x,y) \leqslant N_2} \frac{1}{Q(x, y)},$$

where $Q(x, y)$ runs over the reduced forms with discriminant $-D$.

We easily find from Lemma 4 that

$$\sum_{N_2/D^2 \leqslant Q(x,y) \leqslant N_2} \frac{1}{Q(x,y)} < C_2 \frac{\log_2 D}{\sqrt{D}}.$$

Summation over all forms $Q(x, y) \leqslant 2h(-D)$ yields

$$N_2 \sum_{Q} \sum_{N_2/D_2 \leqslant Q(x,y) \leqslant N_2} \frac{1}{Q(x,y)} < 2N_2 h(-D) \frac{C_2 \log_2 D}{\sqrt{D}}. \qquad (11.3.8)$$

But, by (11.3.3),

$$h(-D) \leqslant D^{1/2-\eta(\beta)},$$

so that we obtain

$$N_2 \sum_{N_2/D^2 \leqslant n' \leqslant N_2} \frac{2^{\nu(n')}}{n'} < 2C_2 N_2 \log_2 D \cdot D^{1/2-\eta(\beta)} \cdot \frac{1}{\sqrt{D}}$$
$$= B_0 N_2 \log_2 D \cdot D^{-\eta(\beta)}. \qquad (11.3.9)$$

We have for sums of the second type, where $p_i p_j \ldots p_\varrho < N_2/D^2$,

$$\Big|2^m \sum_{n \leqslant \frac{N_2}{p_i \cdots p_\varrho}} \chi_k(n)\, \chi'(n)\Big| < N_2^{\varepsilon'} 2\sqrt{kD}\, \ln(kD), \qquad (11.3.10)$$

on the basis of an obvious modification of formula (11.3.6).

Further, $2^m \ll N_2^{\varepsilon'}$ by Lemma 3. The number of sums of the second type is equal to $p_i p_j \ldots p_\varrho < N_2/D^2$, so that their common bound does not exceed

$$C_{\varepsilon'} N_2^{2\varepsilon'} \frac{\sqrt{kD}\, \ln(kD)}{D^2} N^2. \qquad (11.3.11)$$

4. Continuation of the Proof. Further Lemmas

In conjunction with (11.3.7), (11.3.9) and (11.3.11) yield

$$\Big|\sum_{n \leqslant N_2} \chi_k(n)\, \lambda(n)\Big| \ll N_2 D^{-0{\cdot}9\eta(\beta)}, \qquad (11.4.1)$$

while at the same time, by (11.3.4),

$$\Big|\sum_{n \leqslant N_2} \chi_k(n)\, \lambda(n)\Big| > N_2^{\beta-\varepsilon_0} = N_2^{\beta-0{\cdot}01(1-\beta)}. \qquad (11.4.2)$$

In view of the choice of N_2, we have

$$N_2 \leqslant 2D^{\frac{4}{\beta-\frac{1}{2}-\varepsilon_0}} < D^9,$$

whence

$$D^{0\cdot 9\eta(\beta)} > N_2^{0\cdot 1\eta(\beta)}.$$

On recalling that $\eta(\beta) = 10\cdot 5(1-\beta) + 0\cdot 1\eta_0$, we obtain, on comparing (11.4.1) and (11.4.2),

$$N_2^{\frac{1}{2}-1\cdot 01(\beta)} \ll N_2^{\frac{1}{2}-0\cdot 1\eta(\beta)} = N_2^{\frac{1}{2}-1\cdot 05(1-\beta)-0\cdot 1\eta_0}. \tag{11.4.3}$$

When $D > C_0(\beta, \eta_0, k)$ we obtain an obvious contradiction, thereby proving the fundamental Lemma I.

LEMMA 5. *As* $h \to 0$ *we have*

$$\ln(1+h) = 1 + O(h), \quad (1+h)^\beta = 1 + O(h), \tag{11.4.4}$$

where $\beta > 0$ is a fixed number.

This lemma is proved by using the definition of the number e and natural logarithms, together with the properties of the ordinary binomial.

Let $(-k)$ be a fundamental discriminant, while $\chi_k(n) = (-k/n)$ is the corresponding character. Let ε_1 be any fixed number under the conditions

$$0 < \varepsilon_1 < 0\cdot 01. \tag{11.4.5}$$

We assume that

$$h(-k) < k^{1/2-\varepsilon_1}, \tag{11.4.6}$$

and, on the assumption that k is sufficiently large compared with $1/\varepsilon_1$, derive some consequences of (11.4.6).

Let a_1 be a positive number under the conditions

$$0\cdot 01\varepsilon_1 \leqslant a_1 \leqslant \frac{3}{2}. \tag{11.4.7}$$

We shall consider the expression

$$L(a_1) = \sum_{n \leqslant k^3} \chi_k(n)\, n^{-a_1} \tag{11.4.8}$$

and, as we did earlier, attempt to replace the $\chi_k(n)$ in it by $\lambda(n)$, and then pick out the numbers which are sums of squares. In other words, we shall compare (11.4.8) with

$$L_1(a_1) = \sum_{n \leqslant k^3} \lambda(n)\, n^{-a_1}. \tag{11.4.9}$$

Let

$$A = \frac{4}{\varepsilon_1}, \quad (\ln k)^A = \Delta, \quad k_1 = k^3. \tag{11.4.10}$$

Let us examine the result of picking out from (11.4.8) the numbers n containing prime factors not exceeding Δ. If we use the notation

$$\Pi_1(a_1) = \prod_{p<\Delta}\left(1 - \frac{\chi_k(p)}{p^{a_1}}\right) = \sum_q \frac{b_q}{q^{a_1}},$$

we obtain

$$L_2(a_1) = \sum_{q<k_1} \frac{b_q}{q^{a_1}} \sum_{m<k_1/q} \chi_k(m)\, m^{-a_1}, \tag{11.4.11}$$

where $L_2(a_1)$ is the result of sieving out of $L(a_1)$ the above-mentioned numbers.

We shall prove the following: given any $N \geqslant k$, we have

$$\sum_{N<q<2N} |b_q|\, q^{-a_1} \ll N^{-0\cdot15\varepsilon_1}. \tag{11.4.12}$$

For, let (11.4.12) be not satisfied for some $N \geqslant k$. Let $a_1' = 1 - 0\cdot2\varepsilon_1$. We now have

$$\sum_{N<q<2N} |b_q|\, q^{-a_1'} > N^{a_1-a_1'} \sum_{N<q<2N} |b_q|\, q^{-a_1} \gg N^{-0\cdot15\varepsilon_1+0\cdot19\varepsilon_1} = N^{0\cdot04\varepsilon_1}. \tag{11.4.13}$$

On the other hand,

$$\sum_{N<q<2N} \frac{|b_q|}{q^{a_1'}} < \prod_{p<\Delta}\left(1 + \frac{1}{p^{a_1}}\right)$$

and

$$\ln \prod_{p<\Delta}\left(1 + \frac{1}{p^{a_1'}}\right) = B \sum_{n<\Delta} \frac{1}{n^{a_1'}} = B\Delta^{1-a_1'}$$

by Lemmas 1 and 5.

Moreover,

$$\Delta^{1-a_1'} = (\ln k)^{A0\cdot2\varepsilon_1} = (\ln k)^{0\cdot8} \leqslant (\ln N)^{0\cdot8}. \tag{11.4.14}$$

This clearly contradicts (11.4.13) and proves the lemma.

We put

$$\Pi_2(a_1) = \prod_{\substack{\Delta<p<k_1 \\ \chi_k(p)=+1}} \frac{1 - \dfrac{1}{p^{a_1}}}{1 + \dfrac{1}{p^{a_1}}} = 1 + \sum_{q>\Delta} \frac{d_q}{q^{a_1}}, \tag{11.4.15}$$

and prove the following lemma.

LEMMA 6. *Given any $N \in [\Delta, k^4]$ number of primes under the condition $\chi_k(p) = +1$, $N \leqslant p \leqslant 2N$, we have*

$$N_1 \ll N^{-0\cdot015\varepsilon_1}. \tag{11.4.16}$$

Suppose that this is not the case for the given N. We find an integer r, for which $(2N)^r = k^\gamma$ $(2 \leqslant \gamma < 3)$, and consider the $C^r_{N_1}$ possible distinct products in r of our primes. We have

$$C^r_{N_1} > \frac{N_1^r 2^{-r}}{r!} > N_1^r e^{-r(1+\ln r)}$$

(obviously, $r! \leqslant e^{r \ln r}$),

$$r \leqslant \frac{3 \ln k}{\ln \Delta} = \frac{3}{4} \varepsilon_1 \frac{\ln k}{\ln \ln k}, \quad \ln r \leqslant \ln \ln k.$$

Hence we have, with k large,

$$C^r_{N_1} > N_1^r k^{-0.76\varepsilon_1}.$$

If $N_1 > N^{1-0.015\varepsilon_1}$, then

$$C^r_{N^1} > k^\gamma k^{-0.76\varepsilon_1}.$$

All the products constructed lie between k^γ and $k^\gamma 2^{-r}$, $r \leqslant (3\varepsilon_1/4)$ $(\ln k/\ln \ln k)$. All these products are representable by quadratic forms $Q(x, y)$ with the discriminant $(-k)$ and Lemma 4 yields at once in this case

$$h(-k) > k^{1/2-0.82\varepsilon_1},$$

which contradicts (11.4.6) when k is large. The lemma is proved.

On denoting

$$\overline{\Pi}_2(a_1) = \prod_{\substack{\Delta < p \leqslant k_1 \\ \chi_k(p) = +1}} \frac{1 + \dfrac{1}{p^{a_1}}}{1 - \dfrac{1}{p^{a_1}}},$$

we use this lemma to obtain the product majorizing $\Pi_2(a_1)$ in accordance with Lemma 5:

$$\ln \overline{\Pi}_2(a) \ll \prod_{\substack{\Delta < p \leqslant k \\ \chi_k(p) = +1}} \frac{1}{p^{a_1}} \ll \frac{1}{\Delta^{a_1}} \prod_{i=1}^{\infty} \frac{1}{2^{ia_1}}$$

$$\times (\Delta 2^i)^{1-0.015\varepsilon_1} \ll (\ln k)^{-0.02}. \tag{11.4.17}$$

This leads us to

LEMMA 6a. *When* $N \geqslant k$, $a_2 \geqslant 1 - 0.005\varepsilon_1$, *we have*

$$\sum_{N \leqslant q \leqslant 2N} |d_q| q^{-a_2} \ll N^{-0.004\varepsilon_1}. \tag{11.4.18}$$

The proof, using (11.4.17), is similar in all respects to the proof of (11.4.12).

We shall assume that a_3 is fixed and belongs to the interval $[1 - 0{\cdot}001\varepsilon_1, 3/2]$. Putting

$$\Pi(a_3) = \prod_{2\leqslant p\leqslant \Delta} \frac{1-\frac{\chi_k(p)}{p^{a_3}}}{1+\frac{1}{p^{a_3}}} \prod_{\substack{\Delta<p\leqslant k_1\\ \chi_k(p)=+1}} \frac{1-\frac{1}{p^{a_3}}}{1+\frac{1}{p^{a_3}}} \prod_{2\leqslant p\leqslant k_1}\left(1-\frac{1}{p^{2a_3}}\right)$$

$$= \sum_q \frac{e_q}{q^{a_3}},$$

we find, on combining the foregoing formulae, that

$$\ln \Pi(a_3) = B(\ln k)^{0\cdot 8}, \tag{11.4.19}$$

$$\sum_{N\leqslant q\leqslant 2N} |e_q|\, q^{-a_3} = BN^{-0\cdot 003\varepsilon_1} \quad \text{when} \quad N \geqslant k. \tag{11.4.20}$$

We have further:

$$L_1(a_3) = \sum_{n\leqslant k_1} \frac{\mu(n)}{n^{a_3}} \sum_{q\leqslant k_1} \frac{l_q}{q^{a_3}} \sum_{m\leqslant k_1/q} \frac{\chi_k(m)}{m^{a_3}}. \tag{11.4.21}$$

But Abelian summation and an asymptotic formula of the type (11.1.2) yield

$$\sum_{m\leqslant k_1/q} \frac{\chi^k(m)}{m^{a_1}} = \begin{cases} L(a_1) + Bk^{-1/2} & \text{when} \quad q \leqslant k^{1\cdot 5}, \\ B_k^{0\cdot 001\varepsilon_1} & \text{when} \quad q > k^{1\cdot 5}. \end{cases} \tag{11.4.22}$$

We substitute from (11.4.20) and (11.4.22) in (11.4.21), and thus obtain

$$L_1(a_3) = \sum_{q\leqslant k_1/k^2} \frac{l_q}{q^{a_1}} \left(L(a_3) + Bk^{-1/2}\right) + R_{a_1}, \tag{11.4.23}$$

where

$$R_{a_1} \ll k^{0\cdot 0015\varepsilon_1} \sum_{u^2\leqslant q\leqslant k^2} \frac{|l_q|}{q^{a_3}} \ll k^{0\cdot 0015\varepsilon_1 - 0\cdot 003\varepsilon_1} \leqslant k^{-0\cdot 0015\varepsilon_1}. \tag{11.4.24}$$

Further, (11.4.20) yields

$$\sum_{q\leqslant k_1/k^2} \frac{l_q}{q^{a_3}} = \Pi(a_3) + Bk^{-0\cdot 003\varepsilon_1}. \tag{11.4.25}$$

Finally, the familiar asymptotic formula of arithmetic:

$$\tau(n) = O(n^\eta)$$

yields, given any fixed $\eta > 0$,

$$\sum_{q\leqslant k_1/k^2} \frac{|l_q|}{q^{a_1}} \ll k^{0\cdot 0015\varepsilon_1},$$

Consequently, (11.4.24) and (11.4.25) yield

$$L_1(a_3) = \prod(a_3)\, L(a_3) + Bk^{-0.0015\varepsilon_1}. \tag{11.4.26}$$

On taking account of (11.4.19), we obtain from this:

$$L(a_3) = L_1(a_3)\,(\prod(a_3))^{-1} + Bk^{-0.001\varepsilon_1}. \tag{11.4.27}$$

We put $a_3 = \beta$, so that $\beta \in [1 - 0.001\varepsilon_1, 3/2]$.

LEMMA 7. *Given any* $n \in [k^2, k^4]$,

$$\sum_{n \leqslant N} \mu(n) \ll n^{1-0.01\varepsilon_1}. \tag{11.4.28}$$

This lemma is proved by means of a trivial modification to the proof of the first fundamental Lemma: if (11.4.28) were not true, we should arrive at a contradiction to the hypothesis

$$h(-k) < k^{1/2-\varepsilon_1}. \tag{11.4.6}$$

(In essence, Lemma 7 is a particular case of a somewhat modified first fundamental Lemma.)

$$F(\beta, k_2) = \sum_{n \leqslant k_2} a_n n^{-\beta},$$

where

$$a_n = (-1)^{n-1}, \quad k_2 = k_1 k = k^4.$$

In this case we have

$$\sum_{n \leqslant k_2} n^{-\beta} \sum_{\delta \mid n} a_\delta \mu\left(\frac{n}{\delta}\right) = 1 - 2^{1-\beta}; \tag{11.4.29}$$

this can be rewritten as

$$\sum_{q \leqslant k_1 k} \frac{\mu(q)}{q^\beta} \sum_{m \leqslant k_1 k/q} \frac{a^m}{m^\beta} = 1 - 2^{1-\beta}. \tag{11.4.30}$$

Here we can isolate the $q > k_1$ and apply (11.4.28) and the obvious inequality $\left|\sum_{m \leqslant \alpha} a_m\right| \leqslant 1$, while arguing as earlier. We obtain

$$L(\beta)\, F(\beta, k_2) = 1 - 2^{1-\beta} + Bk^{-0.005\varepsilon_1} \tag{11.4.31}$$

or, on taking account of (11.4.27),

$$L(\beta)\, F(\beta, k_2) = (1 - 2^{1-\beta})\,(\prod(\beta)) + Bk^{-0.001\varepsilon_1}, \tag{11.4.32}$$

$$\ln \prod(\beta) \ll (\ln k)^{0.8}. \tag{11.4.33}$$

We obtain from this:

$$\left.\begin{aligned} L(\beta)\, F(\beta, k_2) &> c'_{\varepsilon_1} e^{-c_1 (\ln k)^{0.8}} \quad \text{when} \quad \beta = 1 + 0.001\,\varepsilon_1, \\ L(\beta)\, F(\beta, k_2) &< -c''_{\varepsilon_1} e^{-c_1 (\ln k)^{0.8}} \quad \text{when} \quad \beta = 1 - 0.001\,\varepsilon_1. \end{aligned}\right\} \tag{11.4.34}$$

But

$$F(\beta, k_2) = 1 - \frac{1}{2^\beta} + \frac{1}{3^\beta} - \cdots \pm \frac{1}{k_2^\beta} > 1 - \frac{1}{2^\beta} + \frac{B}{k_2^{1/2}} > \frac{1}{4}.$$

Hence we obtain the extremely important inequalities:

$$\left.\begin{aligned} L(\beta) &> c_{\varepsilon_1} e^{-c_1(\ln k)^{0.8}} \quad \text{when} \quad \beta = 1 + 0{\cdot}001\,\varepsilon_1, \\ L(\beta) &< c_{\varepsilon_1} e^{-c_1(\ln k)^{0.8}} \quad \text{when} \quad \beta = 1 - 0{\cdot}001\,\varepsilon_1. \end{aligned}\right\} \tag{11.4.35}$$

5. The Second Fundamental Lemma. Completion of the Proof

A character $\chi_k(n)$ for which we have

$$h(-k) < k^{1/2-\varepsilon_1}, \tag{11.4.6}$$

must possess the property $\mathfrak{N}(\beta_1)$ with $\beta_1 = 1 - 0{\cdot}08\varepsilon_1$.

We shall prove this by assuming the contrary. Then there exists an $N \geqslant k^{100}$ such that, for any $x \in [\sqrt{N}, N]$, we have

$$\left|\sum_{n \leqslant x} \chi_k(n)\,\mu(n)\right| \leqslant x^{1-0{\cdot}08\varepsilon_1}. \tag{11.5.1}$$

We put $x = N$. The "sieve" method leads to the equation

$$\sum_{q \leqslant N} \chi_k(q)\,\mu(q)\,q^{-\beta} \sum_{m \leqslant N/q} \chi_k(m)\,m^{-\beta} = 1. \tag{11.5.2}$$

We take account of (11.4.22) and the inequality

$$\sum_{N_1 \leqslant m \leqslant N_2} \chi_k(m)\,m^{-\beta} < \frac{k}{N_1^\beta}$$

for any N_1. Now, (11.5.1) and (11.5.2) give us, following the same method as in § 4:

$$M(\beta) \sum_{m \leqslant \sqrt{N}} \frac{\chi_k(m)}{m^\beta} = 1 + BN^{-0{\cdot}05\varepsilon_1}, \tag{11.5.3}$$

where

$$M(\beta) = \sum_{q \leqslant \sqrt{N}} \chi_k(q)\,\mu(q)\,q^{-\beta}.$$

Furthermore, we have

$$\sum_{k_1 \leqslant m \leqslant \sqrt{N}} \chi_k(m)\,m^{-\beta} = Bk^{-1},$$

so that

$$Q(\beta) = \sum_{m \leqslant \sqrt{N}} \frac{\chi_k(m)}{m^\beta} = L(\beta) + Bk^{-1}.$$

Hence we obtain, using (11.4.35):

$$\left.\begin{aligned} Q(\beta) &> \tfrac{1}{2} c_{\varepsilon_1} e^{-c_1 (\ln k)^{0.08}} \quad \text{when} \quad \beta = 1 + 0{\cdot}01\,\varepsilon_1, \\ Q(\beta) &< -\tfrac{1}{2} c^{\varepsilon_1} e^{-c_1 (\ln k)^{0.08}} \quad \text{when} \quad \beta = 1 - 0{\cdot}01\,\varepsilon_1. \end{aligned}\right\} \tag{11.5.4}$$

Since $Q(\beta)$ is a continuous function, there exists a point in the interval $[1 - 0{\cdot}001\varepsilon_1, 1 + 0{\cdot}001\varepsilon_1]$, at which $Q(\beta) = 0$. Putting $\beta = \beta_0$ in the inequality (11.5.3), we arrive at an equality which is absurd for large k; the second fundamental Lemma is thus proved.

We can now prove Siegel's theorem. We take $\varepsilon_1 \in (0, 0{\cdot}01)$. If there is a sufficiently large k for which

$$h(-k) < k^{1/2-\varepsilon_1},$$

then, by the second fundamental Lemma, the character $\chi_k(n)$ possesses the property $\mathfrak{N}(1 - 0{\cdot}08\varepsilon_1)$. But now, by the first fundamental Lemma, we have for sufficiently large D:

$$h(-D) > D^{1/2-\eta(\beta)},$$

$$\eta(\beta) = 10{\cdot}5 \cdot 0{\cdot}08\,\varepsilon_1 + \eta_0 = 0{\cdot}84\,\varepsilon_1 + \eta_0 < 0{\cdot}85\varepsilon_1,$$

since η_0 may be chosen as small as desired.

There is thus only a finite number of the k in question, which is what we had to prove.

CHAPTER 12

THE TRANSCENDENCE OF CERTAIN NUMBER CLASSES

1. Auxiliary Propositions

The transcendence of numbers of the form e^{ω}, where ω is algebraic and $\neq 0$, and of numbers of the form a^b, where $a \neq 0, 1$ is algebraic and b is algebraic irrational, is proved by means of analytic methods. In the case of numbers of the form e^{ω}, the properties of the integral of functions of the form $z^k e^{\omega z}$ are used, while still more subtle analytic properties of the function a^{bz} are used in the case of numbers of the form a^b.

We shall give here a proof of the transcendence of numbers of the form e^{ω} and a^b in the case when ω is real, and when $a > 0$ and b are real, with the above conditions regarding their being algebraic; in the analytic part of this proof, only Rolle's Theorem will be used. We shall require subsequently two well-known lemmas from analysis, the proofs of which we shall give, since they are brief.

LEMMA 1. *If all the numbers $a_{k,s}$ are real and jointly non-zero, and the numbers $\alpha_0, \ldots, \alpha_n$ real and distinct, then the number of real zeros of the function $f_n(x)$:*

$$f_n(x) = \sum_{k=0}^{n} e^{\alpha_k x} \sum_{s=0}^{p_k - 1} a_{k,s} x^s, \quad a_{k,p_k-1} \neq 0 \ (0 \leqslant k \leqslant n), \tag{12.1.1}$$

does not exceed $N_n = \sum_{k=0}^{n} p_k - 1$.

Suppose that the lemma is in fact true for $n < q - 1$. Suppose that the number of zeros of some function $f_q(x)$ is greater than the N_q of the lemma. In this case the number of zeros of the function

$$\frac{d^{p_0}}{dx^{p_0}} e^{-\alpha_0 x} f_q(x) = \sum_{k=1}^{q} e^{(\alpha_k - \alpha_0)x} \sum_{s=0}^{p_k - 1} b_{k,s} x^s = f_{q-1}(x)$$

exceeds the number $N_{q-1} = \sum_{1}^{q} p_k - 1$, by Rolle's theorem. We arrive at a contradiction of our assumption, which is an obvious consequence of the fundamental theorem of algebra. The lemma is proved.

LEMMA 2. *If the number of zeros of a real function $f(x)$ is not less than q in the interval $[a, b]$ and $f(x)$ is continuous together with all its derivatives up to and including the qth, then we have, for $a \leqslant x \leqslant b$,*

$$|f^{(\nu)}(x)| \leqslant \frac{(b-a)^{q-\nu}}{(q-\nu)!} \max_{a \leqslant y \leqslant b} |f^{(q)}(y)| \ (\nu < q). \tag{12.1.2}$$

For, if $f(x)$ has q zeros in (a, b), $f^{(\nu)}(x)$ must have at least $q - \nu$ zeros in (a, b). Let these zeros be $x_1, \ldots, x_{q-\nu}$, each zero being repeated as many times as its multiplicity. Let us now consider the function

$$F(y) = f^{(\nu)}(y) - \lambda \prod_{k=1}^{q-\nu} (y - x_k), \quad \lambda = \frac{f^{(\nu)}(x)}{(x - x_1) \cdots (x - x_{q-\nu})},$$
$$x = x_k (a \leqslant x \leqslant b).$$

This function has $q - \nu + 1$ zeros in (a, b). Hence, by Rolle's Theorem, there exists a ξ $(a \leqslant \xi \leqslant b)$ such that $F^{(q-\nu)}(y)$ has a zero for this ξ in (a, b); in other words, such that

$$F^{(q-\nu)}(\xi) = 0 = f^{(q)}(\xi) - \lambda(q-\nu)!,$$

whence

$$|f^{(\nu)}(x)| = \frac{|f^{(q)}(\xi)|}{(q-\nu)!}(b-a)^{q-\nu} \leqslant \frac{(b-a)^{q-\nu}}{(q-\nu)!} \max_{a \leqslant y \leqslant b} |f^{(q)}(\xi)|.$$

The following is a direct corollary of Lemma 2.

LEMMA 3. *If p and q are sufficiently large integers, $f(x)$ satisfies the conditions*

$$\left.\begin{aligned} f(x) &= \sum_{k=0}^{n} e^{\alpha_k x} \sum_{s=0}^{p_k - 1} a_{k,s} x^s, \ n \leqslant p^2, \ p_k \leqslant p, \\ |a_{k,s}| &\leqslant e^{\theta_0 p}, \ |\alpha_n| \leqslant \theta_0 p, \ 0 \leqslant k \leqslant n, \ 0 \leqslant s \leqslant p, \end{aligned}\right\} \tag{12.1.3}$$

where θ_0 is independent of p, and if $f(x)$ has not less than $q > \theta p^2$ zeros in the interval $(0, \sqrt{p})$, $\theta > 0$ is independent of p and q, then we must have

$$|f^{(\nu)}(x)| < e^{-\frac{1}{2}q \ln p + \gamma_0 q}, \quad \nu < \frac{q}{\ln p}, \quad 0 \leqslant x \leqslant \sqrt{p}, \tag{12.1.4}$$

where γ_0 is independent of p and q.

For we have, by Lemma 2,

$$|f^{(\nu)}(x)| < \frac{p^{q-\nu/2}}{(q-\nu)!} \max |f^{(q)}(x)|$$
$$< e^{\frac{1}{2}q \ln p - 2q \ln p + O(q)} \cdot p^3 e^{\theta q} p^q < e^{-\frac{1}{2}q \ln p + \gamma_0 q},$$
$$\nu < \frac{q}{\ln p}, \quad 0 \leqslant x \leqslant \sqrt{p},$$

where γ_0 is independent of p and q.

We shall now quote a familiar lemma from the theory of algebraic numbers.

LEMMA 4. *If $\alpha_1, \alpha_2, \alpha_3$ are algebraic numbers, A_{k_1, k_2, k_3} are integers, $|A_{k_1, k_2, k_3}| < e^{\gamma_1 p^2}$, where γ_1 is independent of p, then either*

$$|N| = \left| \sum_{k_1=1}^{p^2} \sum_{k_2=1}^{p^2} \sum_{k_3=1}^{p^2} A_{k_1, k_2, k_3} \alpha_1^{k_1} \alpha_2^{k_2} \alpha_3^{k_3} \right| > e^{-\gamma_2 p^2}, \qquad (12.1.5)$$

where γ_2 is independent of p, or $N = 0$.

For, if α_1, α_2, α_3 are algebraic, there must be an integer a such that $a\alpha_1$, $a\alpha_2$, $a\alpha_3$ are integral algebraic numbers of the reduced field R of degree ν. Now, $N_1 = a^{3p^2}N$ will be an integer of this field. If N_1 is not zero, all its conjugates $N_1^{(k)}$ satisfy the inequality $|N_1^{(k)}| < e^{\gamma' p^2}$, where γ' is independent of p, while the number $\prod_1^{\nu} N_1^{(k)}$ will be a rational non-zero integer. Hence follows the inequality of the lemma.

LEMMA 5. *If $a_{k,n}$ $(1 \leqslant k \leqslant q,\ 1 \leqslant n \leqslant p,\ q \geqslant 2p)$ are rational integers, $|a_{k,n}| \leqslant a$, then jointly non-zero rational integers $x_1, \ldots, x_q$ can be found such that*

$$\sum_{k=1}^{q} a_{k,n} x_k = 0 \qquad (n = 1, \ldots, p), \qquad (12.1.6)$$

where the $|x_k|$ $(1 \leqslant k \leqslant q)$ are bounded by the inequalities

$$|x_k| \leqslant x = 3aq. \qquad (12.1.7)$$

For the proof, we consider the linear forms in q variables x_k:

$$L_n = \sum_{1}^{q} a_{k,n} x_k \qquad (1 \leqslant n \leqslant p).$$

We take Cartesian axes in p-dimensional Euclidean space and mark off the values of the forms $L_1, \ldots, L_p$ when the x_k run independently of one another over integral values from the interval $[0, 3qa]$.

All the $(3aq + 1)^q$ points $L_1, \ldots, L_p$ thus obtained fall within the p-dimensional cube $-3a^2q^2 \leqslant y_k \leqslant 3a^2q^2$ of volume $(6q^2a^2)^p$. But, given that $q \geqslant 2p$,

$$(6a^2q^2)^p < (3qa)^q.$$

Thus the number of points obtained, with integral-valued coordinates, is greater than the volume of the p-dimensional cube in which they must lie. In other words, there exist two distinct sets of integers $x_1, \ldots, x_q$ and $x_1', \ldots, x_q'$ such that

$$\sum_{k=1}^{q} a_{k,n} x_k = \sum_{k=1}^{q} a_{k,n} x_k'; \quad \sum_{k=1}^{q} a_{k,n}(x_k - x_k') = 0 \qquad (1 \leqslant n \leqslant p).$$

Since $|x_k - x_k'| < 3aq$, and the numbers $x_k - x_k'$ are jointly non-zero, our lemma is proved.

2. General Theorems on the Transcendence of e^ω and α^β for Algebraic and Real ω, α, β

Our lemmas enable us to prove the following theorems.

THEOREM 12.2.1. *If $\omega \neq 0$ is real and algebraic, then e^ω is not an algebraic number.*

We consider the function

$$f(x) = \sum_{n=0}^{p} \sum_{k=0}^{p} C_{k,n} x^k e^{\omega n x}$$

and suppose that $e^\omega = \omega_1$ is algebraic, along with ω. We take the derivative of order ν of $f(x)$, $\nu \leqslant p^2/\ln p$ at the point $x = s$ $(0 \leqslant s \leqslant \ln^2 p)$. Now,

$$f^{(\nu)}(s) = \sum_{n=0}^{p} \sum_{k=0}^{p} C_{k,n} \sum_{m=0}^{\nu} \frac{\nu!\, k!\, n^{\nu-m} s^{k-m}}{m!(\nu-m)!\,(k-m)!}\, \omega^{\nu-m} e^{\omega n s}.$$

Furthermore, if α is an algebraic integer of degree r, then

$$\alpha^r = \sum_{k=1}^{r} b_k \alpha^{r-k}, \quad \alpha^N = \sum_{k=1}^{r} B_{k,N} \alpha^{r-k}, \tag{12.2.1}$$

where b_k and $B_{k,N}$ are integers.

To determine the $B_{k,N}$, we use the fact that the equation of degree r for α is irreducible, whence it follows that all the numbers $\alpha_2, \alpha_3, \ldots, \alpha_r$ conjugate to $\alpha = \alpha_1$ are distinct, and we have the system of equations for the $B_{k,N}$:

$$\sum_{k=1}^{r} B_{k,N} \alpha_s^{r-k} = \alpha_s^N \qquad (s = 1, \ldots, r).$$

The determinant Δ of this system is a Van der Monde determinant and non-zero. Hence

$$|B_{k,N}| < \left| \sum_{s=1}^{r} C_s \alpha_s^N \right| e^{\gamma N}, \ \gamma > 0, \tag{12.2.2}$$

where γ depends only on α.

Since the numbers ω and $\omega_1 = e^{\omega}$ are algebraic, there is an integer d such that $d\omega$ and $d\omega_1$ are algebraic integers of degrees r_1 and r_2 respectively. We therefore have

$$\begin{aligned} (d\omega)^N &= \sum_{\nu=1}^{r_1} B_{\nu,N}(d\omega)^{r_1-\nu}, \\ (d\omega_1)^N &= \sum_{\nu=1}^{r_2} B'_{\nu,N}(d\omega_1)^{r_2-\nu}, \end{aligned} \tag{12.2.3}$$

where $B_{\nu,N}$ and $B'_{\nu,N}$ are integers and satisfy the inequalities (12.2.2) with the constants $\gamma - \gamma_1$ and γ_2 respectively. On taking into account the greatest possible values of ν and s and using equations (12.2.3), we find that

$$\begin{aligned} d^p f^{(\nu)}(s) = \sum_{\nu_1=0}^{r_1-1} \sum_{\nu_2=0}^{r_2-1} \left[\sum_{k=0}^{p} \sum_{n=0}^{p} C_{k,n} D^{(\nu,s)}_{k,n,\nu_1,\nu_2} \right] (d\omega)^{\nu_1} (d\omega_1)^{\nu_2}, \\ |D^{(\nu,s)}_{k,n,\nu_1,\nu_2}| < e^{\theta p^2}, \end{aligned} \tag{12.2.4}$$

where $D^{(\nu,s)}_{k,n,\nu_1,\nu_2}$ are rational integers, while θ is independent of p for $\nu \leqslant p^2/\ln p$, $s < \ln^2 p$.

By using Lemma 5, we can choose for the $C_{k,n}$ jointly non-zero integers, such as will satisfy the equations

$$\begin{aligned} \sum_{k=0}^{p} \sum_{n=0}^{p} C_{k,n} D^{(\nu,s)}_{k,n,\nu_1,\nu_2} = 0, \quad 0 \leqslant \nu_1 \leqslant r_1 - 1, \quad 0 \leqslant \nu_2 \leqslant r_2 - 1, \\ 0 \leqslant \nu \leqslant \left[\frac{p^2}{\ln p} \right], \quad 0 \leqslant s \leqslant \left[\frac{\ln p}{2\nu_1\nu_2} \right], \quad p > p_0, \end{aligned}$$

where the $C_{k,n}$ also satisfy

$$|C_{k,n}| < e^{2\theta p^2}, \tag{12.2.5}$$

and θ is the constant from inequalities (12.2.4). Having chosen the integers $C_{k,n}$ in this way, we find that $f(x)$ satisfies the conditions

$$\begin{aligned} f(x) = \sum_{k=0}^{p} \sum_{n=0}^{p} C_{k,n} x^k e^{\omega n x}, \quad f^{(\nu)}(s) = 0, \\ |C_{k,n}| \leqslant e^{\theta p^2}, \quad 0 \leqslant \nu \leqslant \frac{p^2}{\ln p}, \quad 0 \leqslant s \leqslant \frac{\ln p}{2\nu_1\nu_2}, \end{aligned}$$

where the $C_{k,n}$ are jointly non-zero integers and θ is independent of p. Thus $f(x)$ has not less than $[p^2/2\nu_1\nu_2]$ zeros in the interval $0 \leqslant x \leqslant \ln^2 p$. Hence $f(x)$ satisfies all the conditions of Lemma 3. Consequently,

$$|f_{(x)}^{(\nu)}| < e^{\frac{p^2 \ln p}{5\nu_1\nu_2} + \gamma_0 p^2}, \quad 0 \leqslant \nu \leqslant \frac{p^2}{p \ln}, \quad 0 \leqslant x \leqslant \ln^2 p. \quad (12.2.6)$$

But the numbers

$$f^{(\nu)}(s), \quad 0 \leqslant \nu \leqslant \left[\frac{p^2}{\ln p}\right], \quad 0 \leqslant s \leqslant [\ln^2 p]$$

are algebraic. By Lemma 4, either $f^{(\nu)}(s) = 0$ or

$$f^{(\nu)}(s) > e^{-\gamma_2 p^2}, \quad (12.2.7)$$

where γ_2 is independent of p, since in our case the conditions of Lemma 4 are satisfied. But when $p > p_1$, inequalities (12.2.6) lead to a contradiction with the inequalities (12.2.7), i.e. $f^{(\nu)}(s) = 0$ $(0 \leqslant \nu \leqslant [p^2/\ln p],\ 0 \leqslant s \leqslant \ln^2 p)$. Hence the number of real zeros of $f(x)$ is certainly greater than $4p^2$, while, by Lemma 1, it cannot be greater than $(p+1)^2$, $p > 2$. This contradiction proves that the numbers ω and e^{ω} cannot be algebraic simultaneously.

THEOREM 12.2.2. *If $a \neq 1$ and $b \neq 1$ are real positive algebraic numbers, then the number $\omega = \ln a/\ln b$, where the real branches of the logarithms are taken, is either rational or transcendental.*

COROLLARY TO THEOREM 12.2.2. *Putting $b = 10$, we find that the common logarithms of all the algebraic numbers are either rational or transcendental.* If we suppose that $b^{\omega} = a$, where ω is algebraic, real, and irrational, and $b \neq 0{\cdot}1$ is positive and algebraic, we arrive at a contradiction with the statement of the theorem. Thus, with the conditions stated above, b^{ω} is always transcendental.

Proof of Theorem 12.2.2. The proof follows exactly the same lines as for the previous theorem.

We again consider the function

$$f(x) = \sum_{k=0}^{p} \sum_{n=0}^{p} C_{k,n} a^{kx} b^{nx},$$

where a and b satisfy the conditions of the theorem. We put $\omega = \ln a/\ln b$ and suppose that ω is algebraic and irrational. Also, let d be an integer such that $d\omega$, da, db are algebraic integers of degrees r_1, r_2, r_3 respectively. Now, by virtue of the representations

(12.2.1) and inequalities (12.2.2), we have

$$\ln^{-\nu} bd p^{\nu} f^{(\nu)}(s) = d p^{\nu} \sum_{k=0}^{p} \sum_{n=0}^{p} C_{k,n}(k + n\omega)^{\nu} a^{ks} b^{ns}$$

$$= \sum_{\nu_1=0}^{r_1-1} \sum_{\nu_2=0}^{r_2-1} \sum_{\nu_3=0}^{r_3-1} \left(\sum_{k=0}^{p} \sum_{n=0}^{p} C_{k,n} D^{(\nu,s)}_{k,n,\nu_1,\nu_2,\nu_3} \right) (d\omega)^{\nu_1} (da)^{\nu_2} (db)^{\nu_3},$$

$$|D^{(\nu,s)}_{k,n,\nu_1,\nu_2,\nu_3}| < e^{\theta p^2}, \quad 0 \leqslant \nu \leqslant \frac{p^2}{\ln p}, \quad 0 \leqslant s \leqslant \ln^2 p. \quad (12.2.8)$$

where all the numbers $D^{(\nu,s)}_{k,n,\nu_1,\nu_2,\nu_3}$ are rational integers, and θ is independent of p.

By Lemma 5, we can take as the $C_{k,n}$ jointly non-zero rational integers such that

$$\sum_{k=0}^{p} \sum_{n=0}^{p} C_{k,n} D^{(\nu,s)}_{k,n,\nu_1,\nu_2,\nu_3} = 0, \quad |C_{k,n}| < e^{2\theta p^2},$$

$$0 \leqslant n \leqslant p, \quad 0 \leqslant k \leqslant p, \quad 0 \leqslant \nu_1 \leqslant r_1 - 1,$$

$$0 \leqslant \nu_2 \leqslant r_2 - 1, \quad 0 \leqslant \nu_3 \leqslant r_3 - 1,$$

$$0 \leqslant s \leqslant \left[\frac{p^2}{\ln p}\right], \quad 0 \leqslant s \leqslant \left[\frac{\ln p}{2 r_1 r_2 r_3}\right], \quad p > p_1,$$

where θ is independent of p and is taken from the inequalities (12.2.8). Having chosen the $C_{k,n}$ in this way, we see that the conditions

$$f^{(\nu)}(s) = 0; \quad 0 \leqslant \nu \leqslant \left[\frac{p^2}{\ln p}\right], \quad 0 \leqslant s \leqslant \left[\frac{\ln p}{2 r_1 r_2 r_3}\right], \quad p > p_1,$$

are satisfied for the function $f(x)$.

Using Lemma 3, the conditions of which are satisfied for our function, we obtain, since in our case $q > p^2/2r_1r_2r_3$:

$$|f^{(\nu)}(s)| < e^{-1/(5 r_1 r_2 r_3) p^2 \ln p}, \quad 0 \leqslant \nu \leqslant \frac{p^2}{\ln p}, \quad 0 \leqslant s \leqslant \ln^2 p, \quad p > p_1. \quad (12.2.9)$$

But, by Lemma 4, the conditions of which are also satisfied, either $f^{(\nu)}(s) = 0$ (s is an integer), or

$$|f^{(\nu)}(s)| > e^{-\gamma p^2}, \quad 0 \leqslant \nu \leqslant \left[\frac{p^2}{\ln p}\right], \quad 0 \leqslant s \leqslant [\ln^2 p]. \quad (12.2.10)$$

But, when $p > p_2$, conditions (12.2.9) and (12.2.10) contradict one another, so that

$$f^{(\nu)}(s) = 0, \quad \nu = 0, 1, \ldots, \left[\frac{p^2}{\ln p}\right], \quad s = 0, 1, \ldots, [\ln^2 p].$$

Thus the number of real zeros of $f(x)$ is not greater than $2p^2$, while at the same time, by Lemma 1, it cannot be greater than $(p + 1)^2$. This contradiction proves the theorem. In the case when ω and a, b are complex algebraic numbers, the simplest proofs are those based on the theory of functions of a complex variable. On adding a further lemma in the domain of real numbers, it would be possible for us to estimate, for example, the deviation of $\ln a/\ln b$ from the rational fraction p/q.

However, we shall not dwell on this.

REFERENCES

1. I. M. VINOGRADOV, *Selected Works* (*Izbrannye trudy*), Akad. Nauk SSSR, Moscow (1952).
2. I. M. VINOGRADOV, The method of trigonometric sums in the theory of numbers, *Trudy MIAN* **23**, 1–111 (1947).
3. I. M. VINOGRADOV, *Fundamentals of the Theory of Numbers* (*Osnovy teorii chisel*), Gost., 1–180, Moscow–Leningrad (1954).
4. I. M. VINOGRADOV, An elementary proof of a general theorem of analytic theory of numbers, *Izv. Akad. Nauk* (6), **19**, 785–795 (1925).
5. I. M. VINOGRADOV, An elementary proof of a theorem of the theory of primes, *Izv. Akad. Nauk.* ser. matem., **17**, 3–12 (1953).
6. G. F. VORONOI, *Collected Works* (*Sobranie sochinenii*) Vol. II, Akad. Nauk. USSR, Kiev (1952).
7. A. O. GEL'FOND, On an arithmetic equivalent of the analyticity of the Dirichlet L-series on the straight line $\mathrm{R}s = 1$, *Izv. Akad. Nauk*, ser. matem., **20**, 145–166 (1956).
8. G. V. EMEL'YANOV, On a system of Diophantine equations, *Uchen. zap. LGU*, ser. matem., **19**, 3–39 (1950).
9. I. P. KUBILYUS, The distribution of primes of a Gaussian field in sectors and contours, *Uchen. zap. LGU*, ser. matem., **19**, 40–52 (1950).
10. I. P. KUBILYUS and YU. V. LINNIK, An elementary theorem of the theory of primes, *Uspek. Matem. Nauk* **11**: 2 (68), 191–192 (1956).
11. YU. V. LINNIK, A note on the least quadratic nonresidue, *Dokl. Akad. Nauk* **36**, 131–132 (1942).
12. YU. V. LINNIK, On Erdös's theorem on the addition of numerical sequences, *Matem. sb.* **10** (52), 67–78 (1942).
13. YU. V. LINNIK, On the representation of large numbers by a sum of seven cubes, *Matem. sb.* **12** (54), 220–224 (1943).
14. YU. V. LINNIK, On the decomposition of large numbers into seven cubes, *Dokl. Akad. Nauk* **36**, 179–180 (1942).
15. YU. V. LINNIK, An elementary proof of Siegel's theorem, based on a method of Vinogradov, *Izv. Akad. Nauk*, ser. matem., **14**, 327–342 (1950).
16. YU. V. LINNIK, An elementary solution of Waring's problem by Schnirelmann's method, *Matem. sb.* **12** (54), 225–230 (1943).
17. YU. V. LINNIK and A. REN'I, On certain hypotheses of the theory of Dirichlet characters, *Izv. Akad. Nauk* **11**, 539–546 (1947).
18. YU. I. MANIN, On congruences of the third degree modulo a prime, *Izv. Akad. Nauk*, ser. matem., **20**, 673–678 (1956).
19. K. K. MARDZHANISHVILI, On the simultaneous representation of n numbers by sums of first, second, ..., nth powers, *Izv. Akad. Nauk*, ser. matem., 609–631 (1937).

20. K. K. MARDZHANISHVILI, On a problem of the additive theory of numbers, *Izv. Akad. Nauk*, ser. matem., 4, 193–214 (1940).
21. B. Z. MOROZ, On the distribution of power residues and nonresidues, *Vestnik LGU*, ser. matem., **19**, 164–169 (1961).
22. A. G. POSTNIKOV and N. P. ROMANOV, A simplification of Selberg's elementary proof of the asymptotic law of distribution of primes, *Uspek. Matem. Nauk*, **10**: 4 (66), 75–87 (1955).
23. N. YA. SONIN, On a definite integral containing the numerical function $[x]$, *Varsh. Univers. izv.* (1885).
24. V. R. FRIEDLÄNDER, On least power nonresidues modulo a prime, *Elabuga, Uchen. zap. ped. in-ta* **1**, 5–55 (1956).
25. I. V. CHULANOVSKII, An elementary proof of the law of distribution of primes of a Gaussian field, *Vestnik LGU* **13**, 43–62 (1956).
26. G. H. HARDY, J. E. LITTLEWOOD and G. POLYA, *Inequalities*, Cambridge University Press, Cambridge (1934).
27. H. HASSE, *Vorlesungen über Zahlentheorie*, Springer, Berlin (1950).
28. A. YA. KHINCHIN, Three pearls of number theory (Tri zhemchuzhiny teorii chisel), *Fizmatgiz* (1948). (Translation published by Graylock Press, Rochester, N.Y., 1952.)
29. A. YA. KHINCHIN, Continued fractions (Tsepnye drobi), *Fizmatgiz* (1949).
30. L. G. SCHNIRELMANN, On the additive properties of numbers, Rostov n/D, *Izv. Dopets. politekhn. in-ta* **14**, 2–3, 3–28 (1930).
31. L. G. SCHNIRELMANN, *Primes* (*Prostye chisla*), Gost., 1–59, Moscow–Leningrad (1940).
32. E. ARTIN and P. SCHERK, On the sum of two sets of integers, *Ann. of Math.* **44** (2), 138–142 (1943).
33. V. BRUN, Le Cridle d'Eratosphene et le théorème de Goldbach, *Videnskaps-selskapet Skr. Mat., naturw.* I, No. 3 (1920).
34. D. BURGESS, The distribution of quadratic residues and nonresidues, *Mathematika* **4** (2), No. 8, 106–112 (1957).
35. H. HASSE, Abstrakte Begründung der komplexen Multiplikation und Riemannsche Vermutung in Funktionenkörpern, *Abh. Math. Sem. Hamburg* **10**, 325–348 (1934).
36. H. DAVENPORT, On the distribution of lth power residues, *Journ. London Math. Soc.* **7** (2), No. 26, 117–121 (1932).
37. H. DAVENPORT, On the distribution of quadratic residues (mod p.), *Journ. London Math. Soc.* **8**, 46–52 (1933).
38. P. ERDÖS, On the arithmetical density of the sum of two sequences one of which forms a basis for the integers, *Acta Arithm., Warszawa*, I, 197–200 (1936).
39. G. H. HARDY, On the expression of a number as the sum of two squares, *Quarterly Journ. Math.* **46**, 263–283 (1915).
40. V. JARNIK, Über Gitterpunkte in mehrdimensionalen Ellipsoiden, *Math. Zeitschr.* 21, 154–160 (1927).
41. E. LANDAU, *Vorlesungen über die Zahlentheorie*, Chelsea, New York (1947).
42. H. B. MANN, A proof of the fundamental theorem on the density of sums of sets of positive integers, *Annals of Math.* **43** (2), 523–527 (1942).

43. S. PILLAI, On Waring's problem, *Journ. Indian Math. Soc.* **2** (2), 213–214 (1937).
44. H. SALIÉ, Über den Kleinsten positiven Nichtrest nach einer Primzahl, *Math. Nachr.* **3**, 7–8 (1949).
45. A. SELBERG, On an elementary method in the theory of primes, *Norske Vid. Selsk. Trondhejm* **19**, No. 18, 64–67 (1947).
46. A. SELBERG, An elementary proof of the prime number theorem, *Ann. Math.* **50**, 305–313 (1949).
47. H. N. SHAPIRO, On primes in arithmetic progression, II, *Ann. Math.* **52**, 231–243 (1950).
48. C. L. SIEGEL, Über die Classenzahl quadratischer Körper, *Acta Arithm.* **1**, 83–86 (1936).
49. I. SIERPINSKI, Sur un probléme du calcul fonctions asymptotiques, *Prace mat.-fis* **17** (1906).
50. A. STOHR and E. WIRSING, Beispiele von wesentlichen Komponenten, die keine Basen sind, *Journ. für die reine und ang. Math.* **196**, 96–98 (1956).
51. A. WEIL, Riemann hypothesis in function fields, *Proc. Mat. Acad. Sc. USA* **27**, 345–347 (1941).